

Old Questions, Big Data: Using the African Digital News (ADN) Corpus to Study News Flows

Dr. Dani Madrid-Morales
University of Houston

Research Seminar | Rhodes University | May 13, 2021

Outline

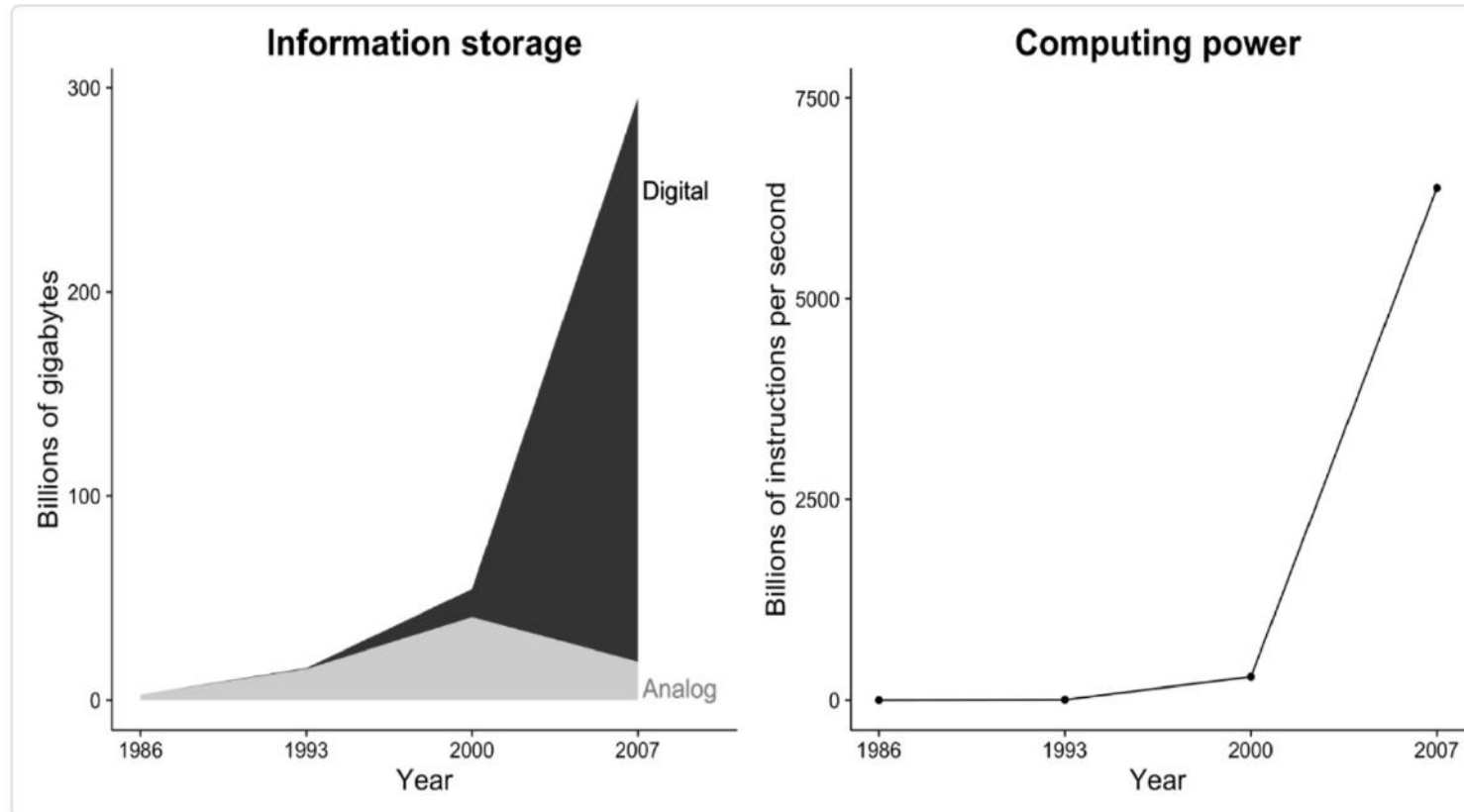
1. Computational Text Analysis
2. African Digital News (adn) corpus
3. Applications
 1. Revisiting the study of News Flows
 2. Revisiting the study of Foreign News
4. Q&A

Computational Text Analysis

Computational Text Analysis

- Social Scientists have always used **texts as data**.
- There are costs (**human labor**) to large-scale text analysis.
- Computers can **lower these costs**.
 - Growth in computational power at relatively low costs.
 - Facilitated by widespread digitization of information.

Computational Text Analysis



Computational Social Sciences

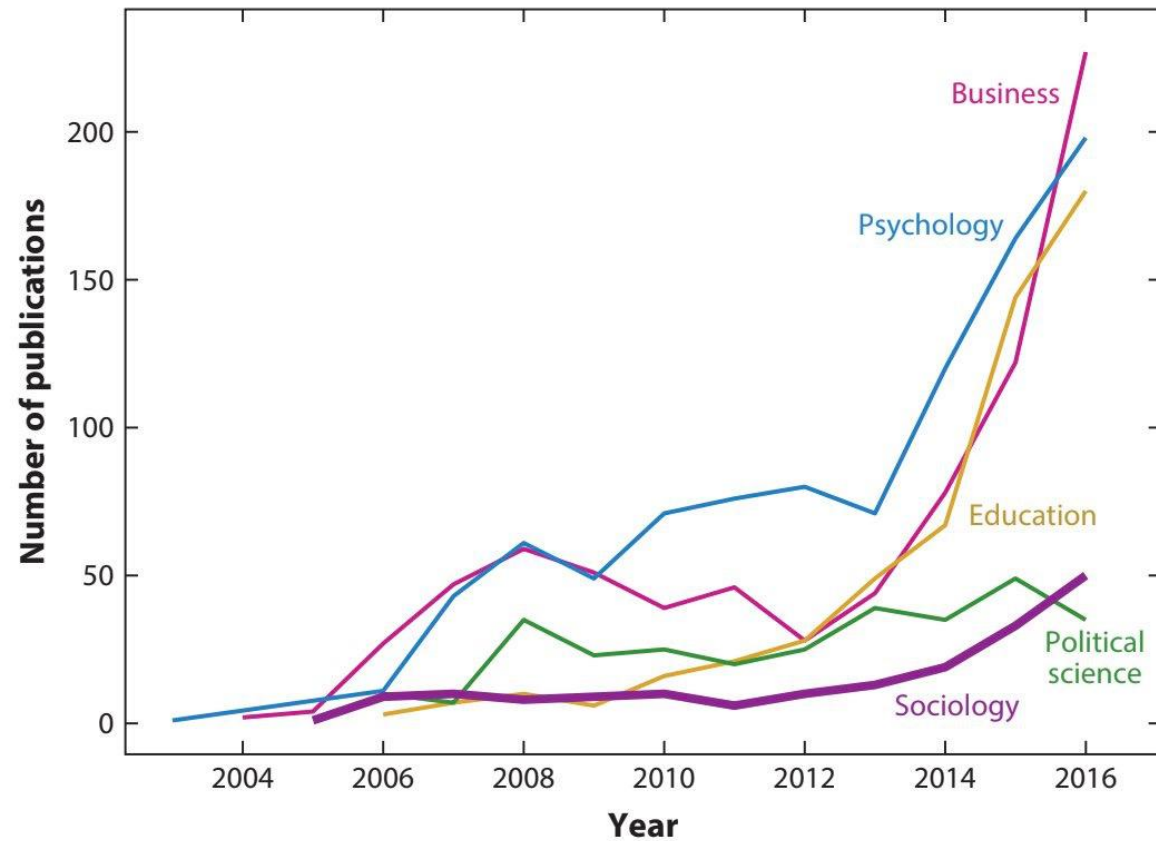


Figure 1

Number of computational social science publications by year—2003–2016—across four scholarly disciplines.

Eddelmann et al. (2020)

Computational Text Analysis

- By using computational tools researchers can...
 - “provide ‘harder’ evidence for what journalism scholars might already have suspected based on qualitative or small-scale quantitative research”
 - “identify patterns in journalistic data that traditional analysis would not”
(Boumans and Trilling 2016, 9)
- In addition, computational approaches make it easier to “revisit” theories, assumptions and questions from **longitudinal** and **comparative** perspective.

Computational Text Analysis – Pros & Cons

- Computational approaches can be useful to answer certain types of questions, but not others.
 - Computational methods are great for **mechanic and repetitive tasks**
 - Computational methods can help **find patterns and classify documents**
 - Computational methods cannot **interpret meanings** of texts
 - Computational methods are **no substitute** of any other research method.
- In other words, computers are good at organizing a haystack, but are terrible (for now) at analyzing individual straws of hay.

African Digital News (adn) corpus

Africa Digital News (adn) Corpus

- In 2019 we started a pilot project to monitor news content on digital news in Africa in real time to create a **live corpus of news**.
- We built this using R and Python.

African Digital News Dataverse (University of Houston)

Texas Data Repository > University of Houston Dataverse Repository > African Digital News Dataverse >

Corpus of African Digital News from 600 Websites Formatted for Text Mining / Computational Text Analysis

Version 3.0

Madrid-Morales, D.; Lindner, P.; Periyasamy, M., 2021, "Corpus of African Digital News from 600 Websites Formatted for Text Mining / Computational Text Analysis", <https://doi.org/10.18738/T8/UKJZ3E>, Texas Data Repository, V3

Cite Dataset ▾

Learn about Data Citation Standards.

Access Dataset ▾

Contact Owner

Share

Dataset Metrics ⓘ

294 Views ⓘ

106 Downloads ⓘ

0 Citations ⓘ

Description ⓘ

This dataset includes a corpus 200,000+ news articles published by 600 African news organizations between December 4, 2020 and January 3, 2021. The texts have been pre-processed (punctuation and English stopwords have been removed, features have been lowercased, lemmatized and POS-tagged) and stored in commonly used formats for text mining/computational text analysis. Users are advised to read the documentation for an explanation of the data collection process.

This dataset includes the following items:

- 31 tables (one per day) of lowercased and lemmatized tokens with the following additional variables: POS tags, document id, sentence id, token id and publication date (stored as a tibble).
- A single document-feature matrix (DFM) with raw counts of feature frequencies in each news article (stored as a quantmod dfm object). The DFM comes with the following metadata for each document: date of publication and source URL.
- A metadata table with the following fields: document id, publication date, source url, news source and country of the news source.
- A list of sources included in the corpus grouped by country name.

All items are stored in formats readable in R. The documentation provides instructions on how to load the RDS files to R.

If you decide to use the data for your own project, please do cite it using the information above. If you identify errors or missing sources, please contact us so that these can be addressed.

Subject ⓘ

Arts and Humanities; Social Sciences

Keyword ⓘ

news, Africa, digital news, text mining, computational

Open Access
Open Source

Algorithmic biases

Google

Tigray

About 1,900,000 results (0.52 seconds)

Yahoo News

Ethiopia's Tigray crisis: 'I lost my hand when a soldier tried to rape me'

A teenager caught up in Ethiopia's Tigray conflict tells of her and her grandfather's near-death encounter.

23 hours ago

Fox News

Ethiopia's secret war in Tigray region: Ethnic killings, rapes, near-starvation reported

Many women have "conclusively and without a doubt" been raped in the Tigray region, home to Ethiopia's secretive conflict - which may have ...

2 days ago

BBC News

Viewpoint: From Ethiopia's Tigray region to Yemen, the dilemma of declaring a famine

Every day, more and more reports of starvation trickle out of the Tigray region of Ethiopia that has been hit by conflict. On Wednesday, Mark ...

1 week ago

Human Rights Watch

Ethiopia: Unlawful Shelling of Tigray Urban Areas

Ethiopian federal forces carried out apparently indiscriminate shelling of urban areas in the Tigray region in November 2020 in violation of the ...

5 days ago

Stay up to date on results for **Tigray**.

Create alert

Google >

1 2 Next

Google

site:addisstandard.com Tigray

When you refine a search with a "site:" operator, Google shows results from all the indexed pages from the website you specified.

About 1,240 results (0.50 seconds)

addisstandard.com

News: Ethiopia "deeply concerned" by EU statement on Tigray, stays silent on accusations of Eritrea's role in the...

News: Ethiopia "deeply concerned" by EU statement on Tigray, stays silent on accusations of Eritrea's role in the armed conflict. addisstandard / ...

6 days ago

addisstandard.com

News Alert: EU joins the US in calling withdrawal of Eritrean troops from Ethiopia, says they are "fueling the..."

EU's latest call on the need to withdraw Eritrean forces from Tigray shows that there is a transatlantic consensus between European Union and ...

1 week ago

addisstandard.com

News: Security forces shot dead Tigray TV journalist, his friend in Mekelle; officials decline to comment

Addis Abeba, January 21/2020 – Dawit Kebede Araya, a journalist working for the regional Tigray TV, currently being administered by the ...

4 weeks ago

addisstandard.com

News Analysis: Ethiopia Rights Commission latest report on Tigray details killings of civilians, widespread looting...

EHRC further stated that the findings of the Commission's on-going investigation missions in Mekelle other parts of Tigray region since January ...

4 weeks ago

Stay up to date on results for **site:addisstandard.com Tigray**.

Create alert

Google >

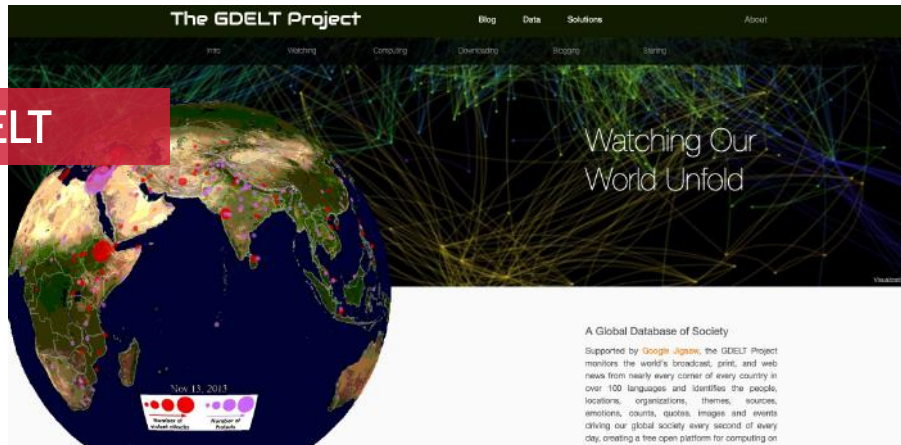
1 2 3 Next

Limitations of existing archives/corpora

Nexis Uni

NewsAPI

GDELT



NOW Corpus

Content Analysis of African News Media

Table 1. Academic articles using content analysis to study news content in Africa published between 2015 and 2019, in selected journals.

	Percentage (N = 47)
Object of study	
Newspapers	40.4
Social media	23.4
News websites	21.3
TV	4.3
Magazines	4.3
Other	6.3
Region of study	
Multiple countries	14.9
Single country—Central Africa	2.1
Single country—East Africa	12.8
Single country—North Africa	4.3
Single country—Southern Africa	44.6
Single country—West Africa	21.3
Sampling method	
Not random	80.8
Random	2.1
Census	12.8
Not specified	4.3
Sample size	
Under 50	25.5
Between 51 and 200	25.5
Between 201 and 500	14.9
Between 501 and 1000	10.6
Over 1001	17.1
Not specified	6.4
Data collection method	
Manual retrieval	63.8
Website search	12.9
Data mining	8.5
Database search	2.1
Google search	2.1
Not specified	10.6

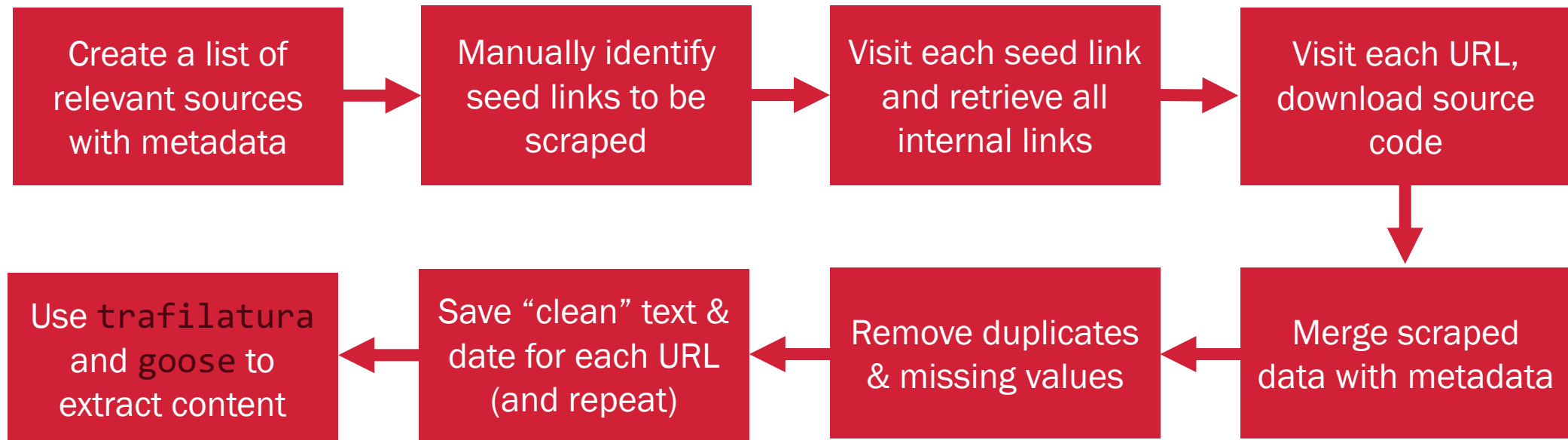
Relatively limited research on online news content

Very few studies are comparative

Sample size of most studies is small (less than 1,000)

Data collection is often manual (i.e. labor intensive)

adn Corpus



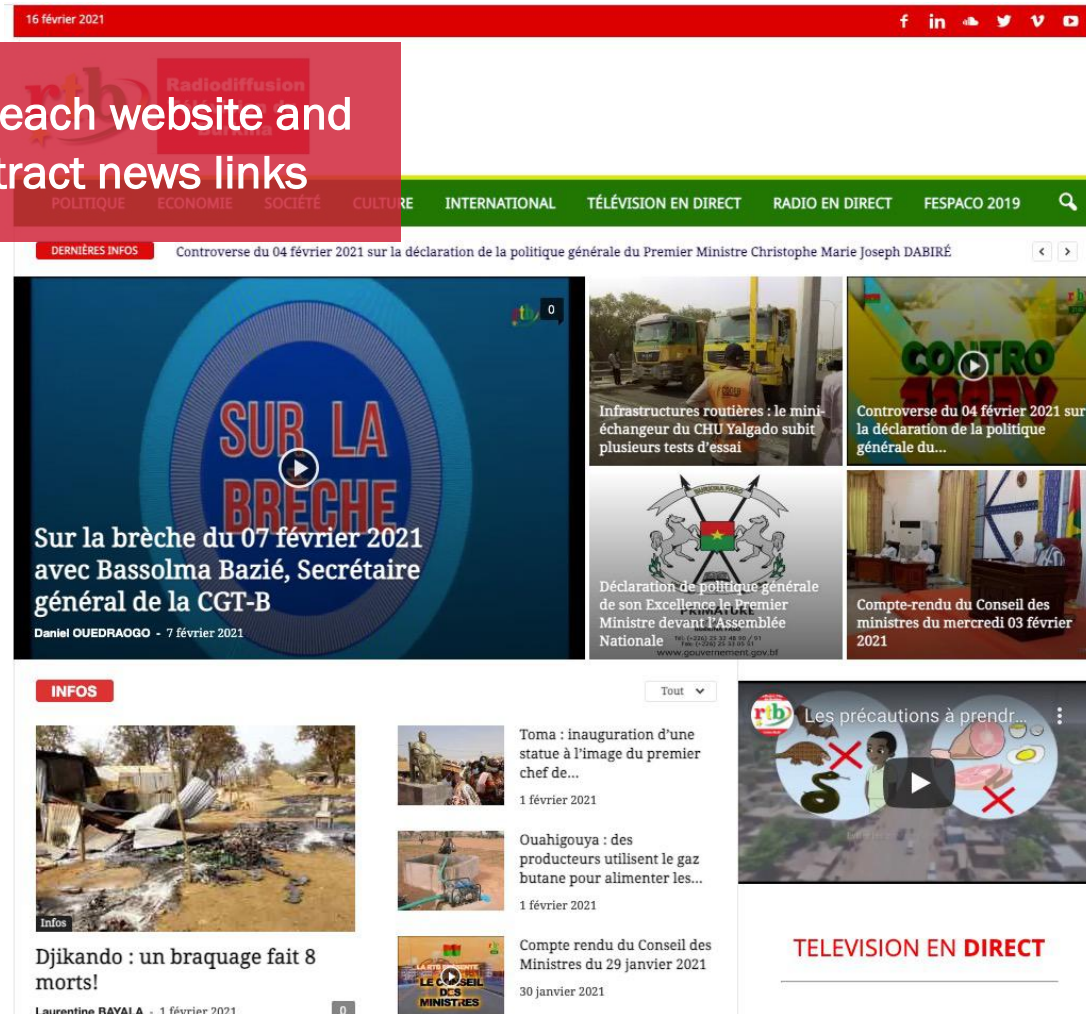
Process – Step 1: Seed Links

Create a list of pages
you want to scrape

A	B	C	D	E	F	G
country_name	country_is	medium_name	alexa_sco	source	url	news_site
Burkina Faso	BF	http://www.aib.media	NA	xyzlinks	http://www.aib.media	TRUE
Burkina Faso	BF	http://www.aouaga.com	NA	xyzlinks	http://www.aouaga.com	TRUE
Burkina Faso	BF	http://www.aujourd8.net	NA	xyzlinks	http://www.aujourd8.net	TRUE
Burkina Faso	BF	http://www.burkina24.com	NA	xyzlinks	http://www.burkina24.com	TRUE
Burkina Faso	BF	http://www.burkinainfo.com	NA	xyzlinks	http://www.burkinainfo.com	TRUE
Burkina Faso	BF	http://www.faso-actu.net	NA	xyzlinks	http://www.faso-actu.net	TRUE
Burkina Faso	BF	http://www.fasopresse.net	NA	xyzlinks	http://www.fasopresse.net	TRUE
Burkina Faso	BF	http://www.fasozine.com	NA	xyzlinks	http://www.fasozine.com	TRUE
Burkina Faso	BF	http://www.leconomistedufaso.bf	NA	xyzlinks	http://www.leconomistedufaso.bf	TRUE
Burkina Faso	BF	http://www.lefaso.net	NA	xyzlinks	http://www.lefaso.net	TRUE
Burkina Faso	BF	http://www.lepays.bf	NA	xyzlinks	http://www.lepays.bf	TRUE
Burkina Faso	BF	http://www.lepays.bf	NA	xyzlinks	http://www.lepays.bf	TRUE
Burkina Faso	BF	http://www.lesaffairesbf.com	NA	xyzlinks	http://www.lesaffairesbf.com	TRUE
Burkina Faso	BF	http://www.lexpressdufaso-bf.com	NA	xyzlinks	http://www.lexpressdufaso-bf.com	TRUE
Burkina Faso	BF	http://www.lobserveur.bf	NA	xyzlinks	http://www.lobserveur.bf	TRUE
Burkina Faso	BF	http://www.lobserveur.bf	NA	xyzlinks	http://www.lobserveur.bf	TRUE
Burkina Faso	BF	http://www.monpulsar.com	NA	xyzlinks	http://www.monpulsar.com	TRUE
Burkina Faso	BF	http://www.mutationsbf.info	NA	xyzlinks	http://www.mutationsbf.info	TRUE
Burkina Faso	BF	http://www.omegabf.info	NA	xyzlinks	http://www.omegabf.info	TRUE
Burkina Faso	BF	http://www.ouagafm.bf	NA	xyzlinks	http://www.ouagafm.bf	TRUE
Burkina Faso	BF	http://www.radiob24.com	NA	xyzlinks	http://www.radiob24.com	TRUE
Burkina Faso	BF	http://www.rtb.bf	NA	xyzlinks	http://www.rtb.bf	TRUE
Burkina Faso	BF	http://www.savane fm.bf	NA	xyzlinks	http://www.savane fm.bf	TRUE
Burkina Faso	BF	http://www.sidwaya.info	NA	xyzlinks	http://www.sidwaya.info	TRUE
Burkina Faso	BF	http://www.sidwaya.info	NA	xyzlinks	http://www.sidwaya.info	TRUE
Burundi	BI	http://www.abpinfos.com	NA	xyzlinks	http://www.abpinfos.com	TRUE
Burundi	BI	http://www.arib.info	NA	xyzlinks	http://www.arib.info	TRUE
Burundi	BI	http://www.burundi24.wordpress.com	NA	xyzlinks	http://www.burundi24.wordpress.com	TRUE
Burundi	BI	http://www.burundi-agnews.org	NA	xyzlinks	http://www.burundi-agnews.org	TRUE
Burundi	BI	http://www.burundi-eco.com	NA	xyzlinks	http://www.burundi-eco.com	TRUE
Burundi	BI	http://www.burundi-forum.org	NA	xyzlinks	http://www.burundi-forum.org	TRUE
Burundi	BI	http://www.burundi-information.net	NA	xyzlinks	http://www.burundi-information.net	TRUE
Burundi	BI	http://www.fr.itaraburundi.com	NA	xyzlinks	http://www.fr.itaraburundi.com	TRUE
Burundi	BI	http://www.isanganiro.org	NA	xyzlinks	http://www.isanganiro.org	TRUE
Burundi	BI	http://www.itaraburundi.com	NA	xyzlinks	http://www.itaraburundi.com	TRUE
Burundi	BI	http://www.iwacu-burundi.org	NA	xyzlinks	http://www.iwacu-burundi.org	TRUE

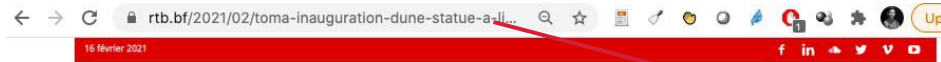
Process – Step 2: Retrieve links

Visit each website and
extract news links



Link	URL
Toma : inauguration d'une statue à l'image du premier chef de...	https://www.rtb.bf/2021/02/toma-inauguration-dune-statue-a-limage-du-premier-chef-de-canton/
Ouahigouya : des producteurs utilisent le gaz butane pour alimenter les...	https://www.rtb.bf/2021/02/ouahigouya-des-producteurs-utilisent-le-gaz-butane-pour-alimenter-les-motopompes/
Compte rendu du Conseil des Ministres du 29 janvier 2021	https://www.rtb.bf/2021/01/compte-rendu-du-conseil-des-ministres-du-29-janvier-2021/
Compte-rendu du conseil des ministres du vendredi 29 janvier 2021	https://www.rtb.bf/2021/01/compte-rendu-du-conseil-des-ministres-du-vendredi-29-janvier-2021/

Process – Step 3: Extracting data



Extract data that is relevant for us

Toma : inauguration d'une statue à l'image du premier chef de canton.



Image MCAT

rtb TOUJOURS AU COEUR DES GRANDS ÉVÈNEMENTS

Les autorités municipales de Toma ont procédé ce samedi 30 janvier 2021 au baptême d'un rond-point où trône une statue d'Issa Gue Paré. Cette statue immortalise le tout premier chef de canton de Toma de 1916 à 1943. Pendant 27 ans de règne, Issa Gue Paré a consacré sa vie au rayonnement de son canton dans plusieurs domaines

publish_date	source	country_iso	full_text	web_url
2020-04-07	abidjan.net	CI	Société Hamed Bakayoko déclaré positif au coronaviru...	https://news.abidjan.net/h/671218.html
2020-04-07	abidjan.net	CI	Société COVID 19: Un don en vivres et non vivres offe...	https://news.abidjan.net/h/672003.html
2020-04-20	abidjan.net	CI	Économie Covid-19: le directeur général de IOMS peu...	https://news.abidjan.net/h/671694.html
2020-04-22	abidjan.net	CI	Économie Hausse du coût de la viande : les acteurs in...	https://news.abidjan.net/h/671803.html
2020-04-29	abidjan.net	CI	Société Secteur du tourisme et des loisirs : les résulta...	https://news.abidjan.net/h/672090.html
2020-03-21	abidjan.net	CI	Économie Covid-19: Les mesures urgentes prises par ...	https://news.abidjan.net/h/670625.html
2020-03-23	abidjan.net	CI	Société Journée mondiale de l'eau 2020 toujours d'ac...	https://news.abidjan.net/h/670673.html
2020-03-27	abidjan.net	CI	International New York, nouvel épice de la pandé...	https://news.abidjan.net/h/670803.html
2020-05-06	abidjan.net	CI	Société Ramadan : Hamed Bakayoko offre des vivres a...	https://news.abidjan.net/h/672361.html
2020-05-12	abidjan.net	CI	Société Coronavirus : Le Premier Ministre met la minis...	https://news.abidjan.net/h/672617.html
2020-05-14	abidjan.net	CI	Société Reprise des travaux de la Nationale A3 à Nlak...	https://news.abidjan.net/h/672735.html
2020-05-27	abidjan.net	CI	Politique Présidentielle 2020: la révision de la liste éle...	https://news.abidjan.net/h/673528.html
2020-02-27	abidjan.net	CI		https://news.abidjan.net/h/669714.html
2020-04-03	abidjan.net	CI	Économie Covid- 19 Des producteurs de café-cacao r...	https://news.abidjan.net/h/671048.html
2020-04-05	abidjan.net	CI	Société Le préfet de Korhogo appelle la population à s...	https://news.abidjan.net/h/671111.html
2020-04-10	abidjan.net	CI	Région COVID-19 : Agboville lance la phase répressiv...	https://news.abidjan.net/h/671331.html
2020-04-12	abidjan.net	CI	Société Pâques 2020 : L'Eglise méthodiste plaide pour ...	https://news.abidjan.net/h/671409.html
2020-04-15	abidjan.net	CI	Art et Culture Covid 19: Des artistes et journalistes se...	https://news.abidjan.net/h/671480.html
2020-04-10	abidjan.net	CI	Société Coronavirus : Le centre de dépistage de Yopo...	https://news.abidjan.net/h/671348.html
2020-04-04	abidjan.net	CI	Société Covid-19 : Sensibilisation à outrance à l'intéri...	https://news.abidjan.net/h/671076.html
2020-04-09	abidjan.net	CI	Société Covid-19: Voici les techniques utilisées par d...	https://news.abidjan.net/h/671283.html
2020-04-03	abidjan.net	CI	Société Treichville/ Lutte contre le Coronavirus : le co...	https://news.abidjan.net/h/671067.html
2020-04-15	abidjan.net	CI	Société Covid-19: l'État de Côte d'Ivoire proroge le dél...	https://news.abidjan.net/h/671500.html
2020-04-19	abidjan.net	CI	Société Bruno Koné fait des dons aux ressortissants d...	https://news.abidjan.net/h/671672.html
2020-03-06	abidjan.net	CI	Société Social: Report du gala de la Fondation Childre...	https://news.abidjan.net/h/670091.html
2020-03-18	abidjan.net	CI	Société Coronavirus: Latelier de formation des format...	https://news.abidjan.net/h/670503.html
2020-03-19	abidjan.net	CI	Politique Covid-19: La plateforme pro-Gbagbo, EDS, l...	https://news.abidjan.net/h/670561.html
2020-03-17	abidjan.net	CI	Société Crise du Coronavirus : Les Abidjanais vident l...	https://news.abidjan.net/h/670484.html
2020-03-27	abidjan.net	CI	Économie Impacts du Covid-19 dans l'UEMOA : Avis r...	https://news.abidjan.net/h/670824.html
2020-03-26	abidjan.net	CI	Économie Covid-19/Abidjan : Le marché gourou ferme...	https://news.abidjan.net/h/670770.html
2020-05-06	abidjan.net	CI	Société Lutte contre le Covid-19: démarrage officiel d...	https://news.abidjan.net/h/672345.html
2020-05-04	abidjan.net	CI	Société Ramadan : 400 familles démunies soulagées à ...	https://news.abidjan.net/h/672295.html
2020-05-10	abidjan.net	CI	Société Lutte contre le Covid-19: les commerçants du...	https://news.abidjan.net/h/672540.html
2020-05-13	abidjan.net	CI	Société Des enseignants musulmans offrent des kits a...	https://news.abidjan.net/h/672674.html
2020-05-12	abidjan.net	CI	Société Covid-19: des équipements sanitaires offerts ...	https://news.abidjan.net/h/672628.html
2020-05-20	abidjan.net	CI	Société Nuit du Qadr: Les musulmans ivoiriens invités...	https://news.abidjan.net/h/673090.html

What does the code look like

```
##Function: Grab links from initial URLs##
get_links <- function(url){

  withTimeout({Sys.sleep(2)

    page <- read_html(url)

    web_link <- page %>%
      html_nodes("a") %>%
      html_attr("href")

    text_links <- page %>%
      html_nodes("a") %>%
      html_text()

    tibble(web_link,
           text_links,
           url)
  }, timeout = 10, onTimeout = "silent")
}
###
```

```
##Function: scrapes out text, date, and raw HTML data from a given link##
get_text_from_py <- function(url_link){

  withTimeout({
    traf <- import("trafilatura")
    requests <- import("requests")
    rand <- import("requests_random_user_agent")
    goose <- import("goose3")

    r <- traf$fetch_url(url_link)
    web_text <- traf$extract(r)
    web_raw1 <- r[1]

    web_date <- traf$metadata$extract_metadata(r) %>%
      str_extract("[0-9]{4}-[0-9]{2}-[0-9]{2}")

    g = goose$Goose({browser_user_agent = browser_agent})

    article <- g$extract(url = url_link)
    web_text2 <- article$cleaned_text
    web_date2 <- as.character(article$publish_datetime_utc)
    web_raw2 <- article$raw_html

    if (is.null(web_date2) == TRUE){
      web_date2 <- ""
    }

    if (is.null(web_text2) == TRUE){
      web_text2 <- ""
    }

    g$close()

    tibble(web_url = url_link,
           web_date = web_date,
           web_date2 = web_date2,
           web_sysdate = Sys.Date(),
           web_text = web_text,
           web_text2 = web_text2,
           web_raw1 = web_raw1,
           web_raw2 = web_raw2)
  }, timeout = 10, onTimeout = "silent")
}
```

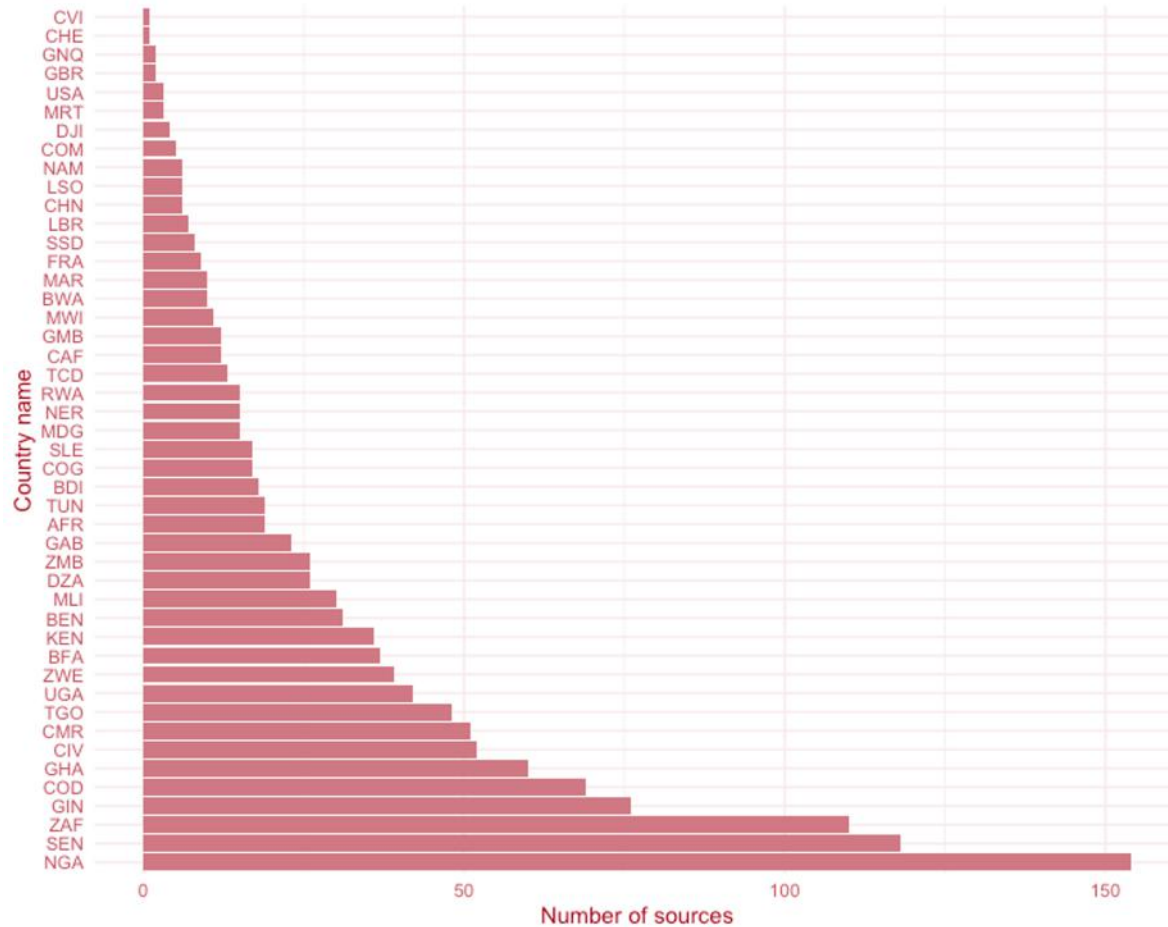
Step 3: Processing the data for analysis

- Use spacy (or udpipe) to tokenize (or lemmatize) the text and POS (part of speech tag)
- We can then remove tokens that we don't "need" (stopwords, numbers, punctuation...)

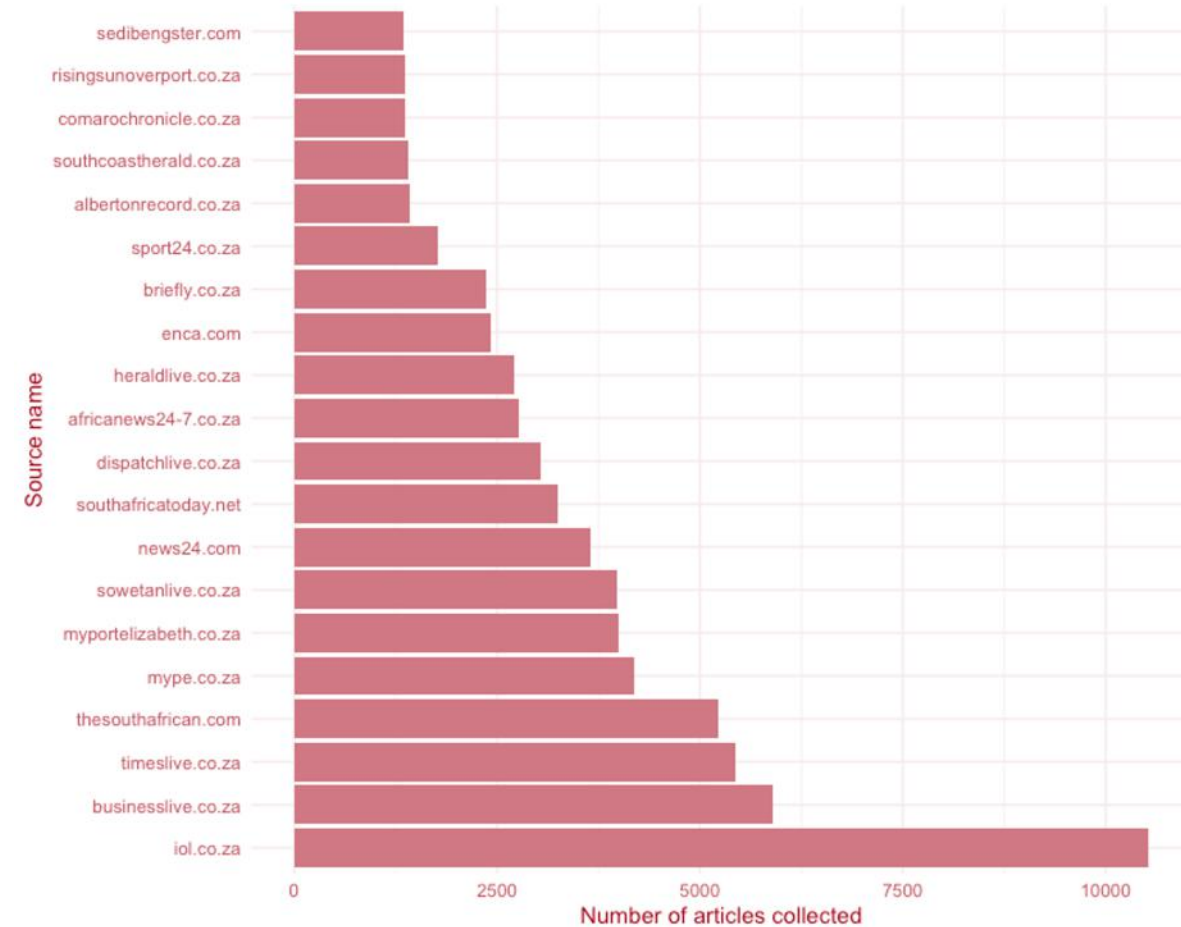
	doc_id	sentence_id	token_id	token	lemma	pos	entity
1	STMID_006086	1	1	When	when	ADV	
2	STMID_006086	1	2	Kem	Kem	PROPN	
3	STMID_006086	1	3	Senou	Senou	PROPN	
4	STMID_006086	1	4	Pavel	Pavel	PROPN	
5	STMID_006086	1	5	Daryl	Daryl	PROPN	
6	STMID_006086	1	6	,	,	PUNCT	
7	STMID_006086	1	7	a	a	DET	
8	STMID_006086	1	8	21-year	21-year	NUM	DATE_B
9	STMID_006086	1	9	-	-	PUNCT	DATE_I
10	STMID_006086	1	10	old	old	ADJ	DATE_I
11	STMID_006086	1	11	Cameroonian	cameroonian	ADJ	NORP_B
12	STMID_006086	1	12	student	student	NOUN	
13	STMID_006086	1	13	living	live	VERB	
14	STMID_006086	1	14	in	in	ADP	
15	STMID_006086	1	15	the	the	DET	
16	STMID_006086	1	16	Chinese	chinese	ADJ	NORP_B
17	STMID_006086	1	17	city	city	NOUN	
18	STMID_006086	1	18	of	of	ADP	
19	STMID_006086	1	19	Jingzhou	Jingzhou	PROPN	GPE_B
20	STMID_006086	1	20	,	,	PUNCT	

adn Corpus | Content

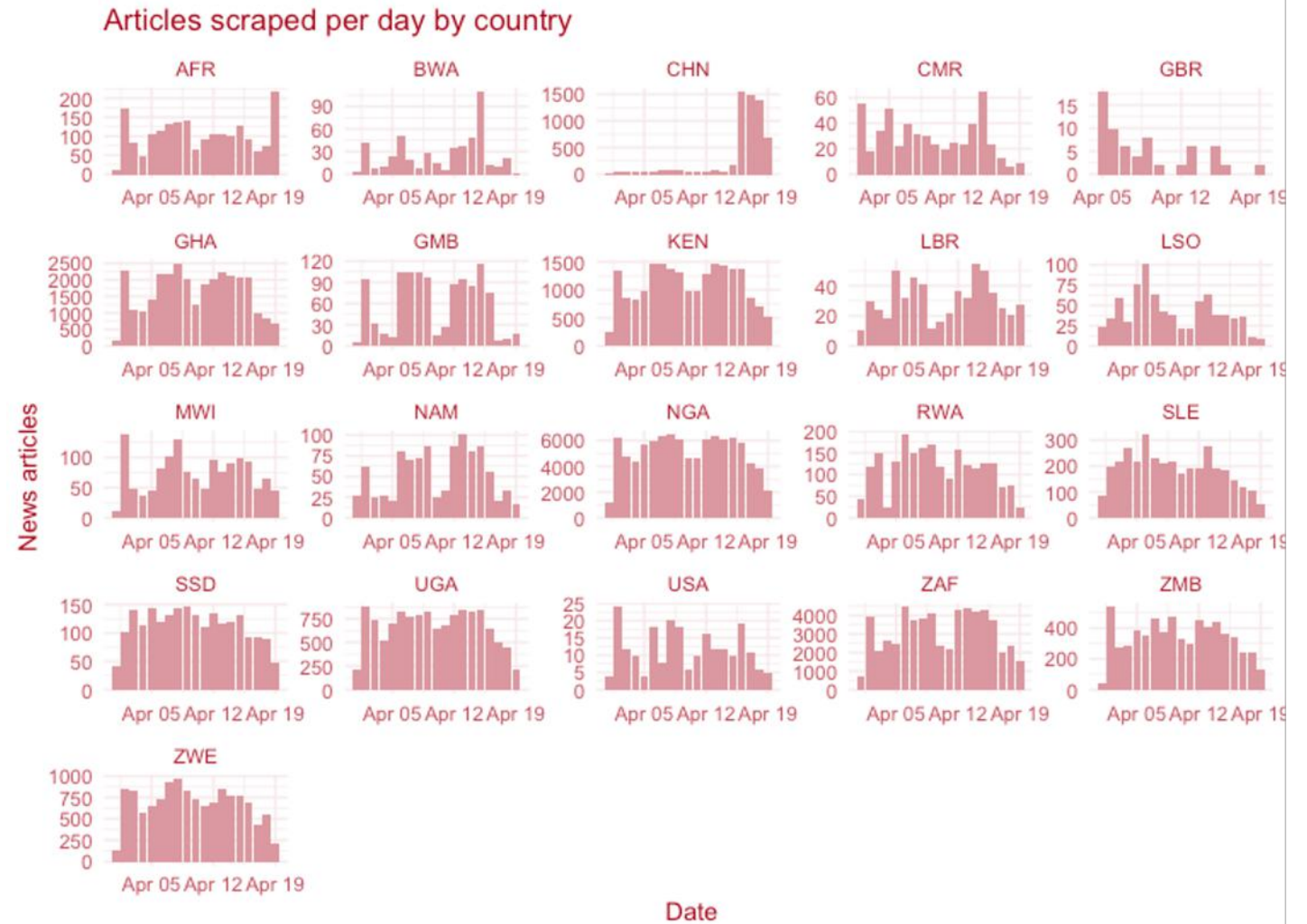
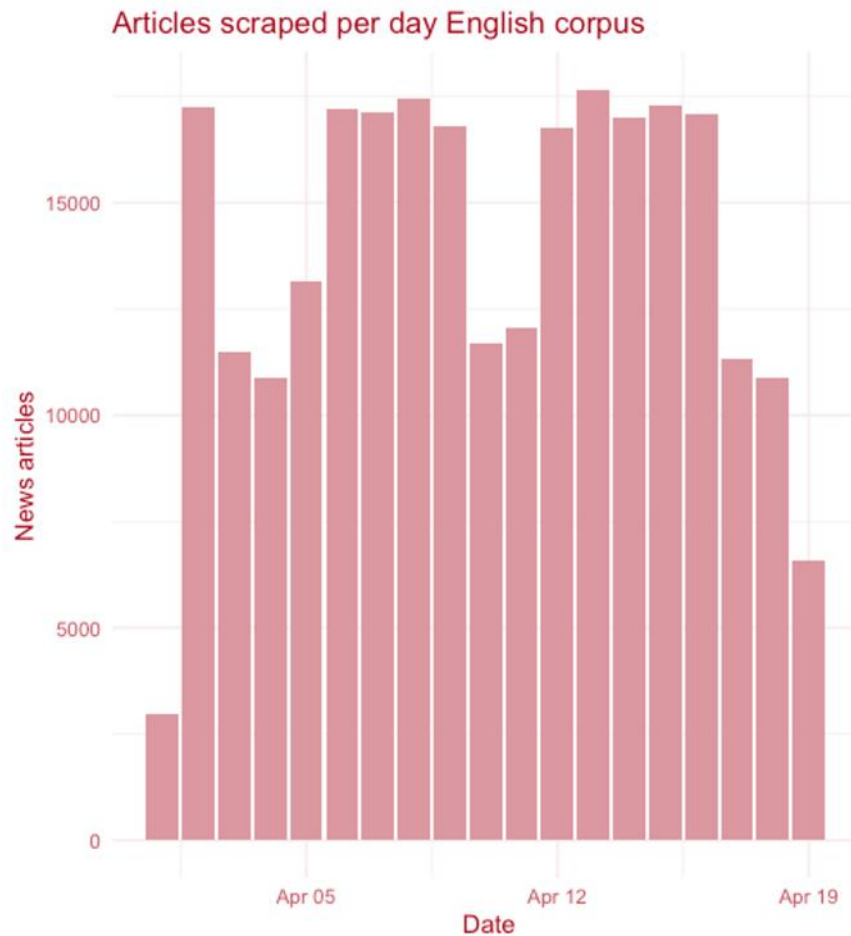
Sources included in the corpus (N = 1,294)



Top 20 South African sources by items scraped



adn Corpus | Content



Limitations

- This approach would (in principle) exclude data that is behind **paywalls**; it is internet data heavy.
- These approaches are heavy on the **amount of bandwidth** that is required (cloud computing as alternative).
- Downtime of websites and other technical problems might limit how much data we can collect, so it is important to acknowledge **biases in our dataset**.

Applications

- With large corpora of textual data...
 - We can answer **descriptive questions** about content
 - We can answer questions about **linguistic patterns**
 - We can **compare sources** across geographic spaces
 - We can trace **longitudinal patterns**
 - We can **triangulate text data with fieldwork**

More information

- There's more details on the process in a paper that came out last year

Madrid-Morales, D. (2020). Using Computational Text Analysis Tools to Study African Online News Content. *African Journalism Studies*, 41(4), 68–82. <https://doi.org/10.1080/23743670.2020.1820885>

- Data from a pilot study of the adn corpus can be downloaded from Madrid-Morales, D., Lindner, P., & Periyasamy, M. (2021). *Corpus of African Digital News from 600 Websites Formatted for Text Mining / Computational Text Analysis* [Data set]. Texas Data Repository. <https://doi.org/10.18738/T8/UKJZ3E>

Revisiting News Flows

Global News Flows

- The **impact that global news agencies** from the Global North have on the news content of media in the Global South has long been a preoccupation of scholars of global communication since Cold War.
- Spurred by the arrival of Qatar's Al Jazeera in the 2000s, there's been an increasing interest in the **possible impact of State-owned media on global news narratives**.
 - These State-owned/backed media include China's and Russia's, but also France's and others'.
 - Chinese media's **actual influence** on African media is poorly understood.

Intermedia agenda setting

- The ability of **well-established news organizations** to have an impact on the content of smaller or **less-established news organizations** (McCombs, 2004).
- Guo and Vargo (2020, p. 3) use the term “**transnational intermedia agenda setting**” to refer to
 - the ability of news media in a country to “transfer the salience of issues associated with their own country to foreign media, thus influencing how foreign media portray the country”
 - “the capability of a country’s news media to push their perspectives on international politics outside of their own country to the world”.

Background

Source: <https://www.chinafile.com/china-africa-project/chinas-media-challenges-western-narratives-of-africa>



Increased presence
of Chinese media in
Africa

← Tweet

Chinese Ambassador to Zimbabwe Retweeted

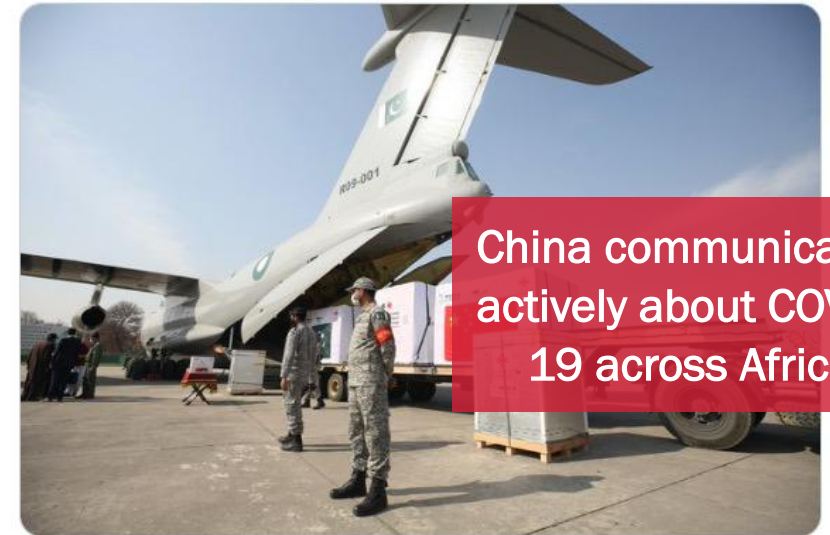


China Xinhua News

@XHNews

China state-affiliated media

China is providing vaccine aid to 53 developing countries: FM xhne.ws/HUSy6



China communicated
actively about COVID-
19 across Africa

4:00 AM · Feb 8, 2021 · China Xinhua News

73 Retweets 13 Quote Tweets 285 Likes

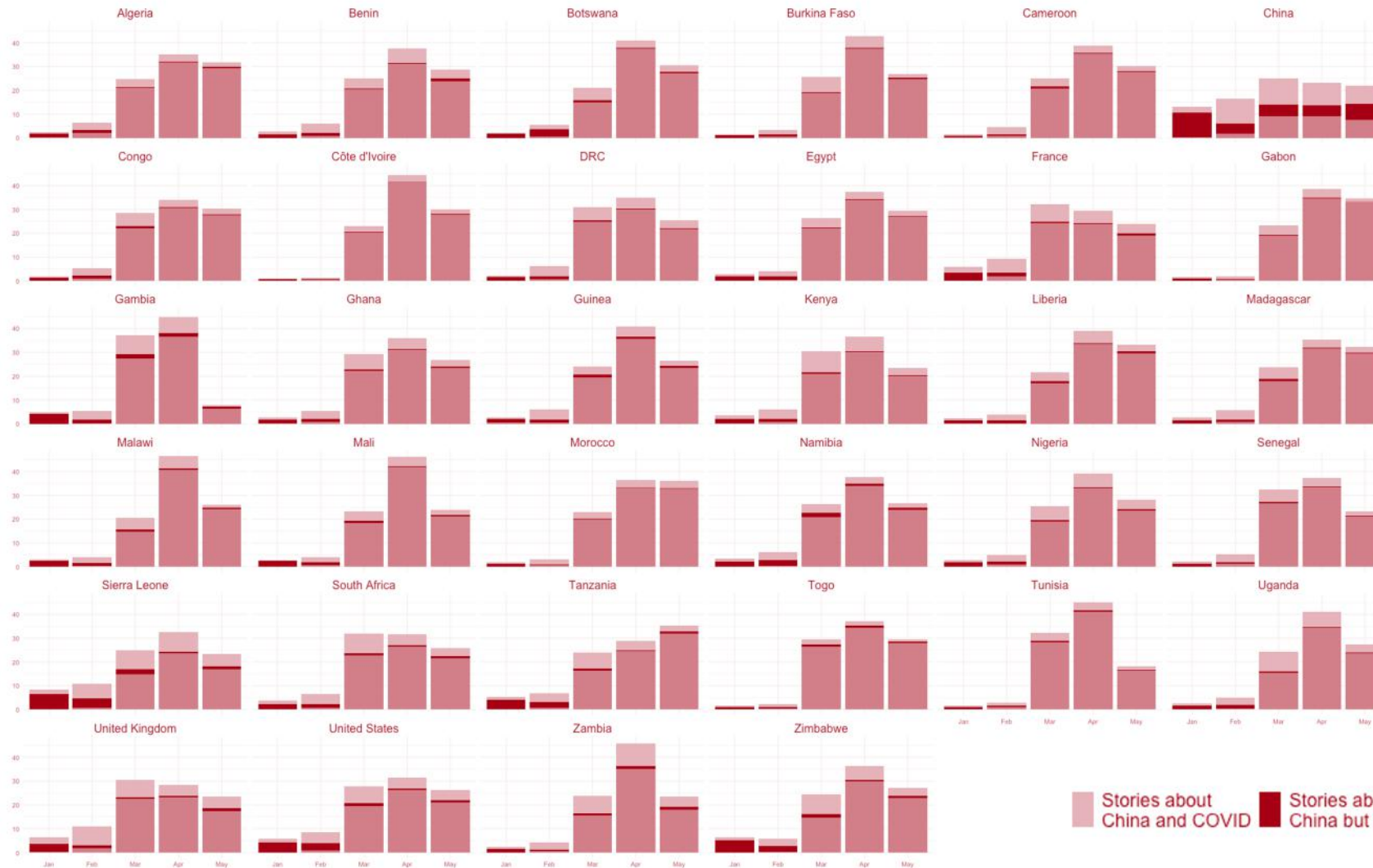
Research Questions

- RQ1: How prevalent were news stories about China and COVID-19 in African media between January and May 2020?
- RQ2: How much content from global media did African news organizations reuse in their reporting of the COVID-19 outbreak?
- RQ3: To what extent was the reuse of Chinese media's content different in Francophone and Anglophone Africa?
- RQ4: How much of China's narrative during the crisis was reflected on African media's coverage of the pandemic?

Analysis

- Compiled 500,000 news stories from French & English language **digital sources based in Africa (16 French- and 16 English-speaking) & UK, France, US, China.**
- Approach 1: Use **word frequencies** to compare texts and find how similar they are
 - RNewsFlow
 - Cosine similarity > .97
- Approach 2: Use word frequencies (and metadata) to **classify documents** into N topics.
 - stm

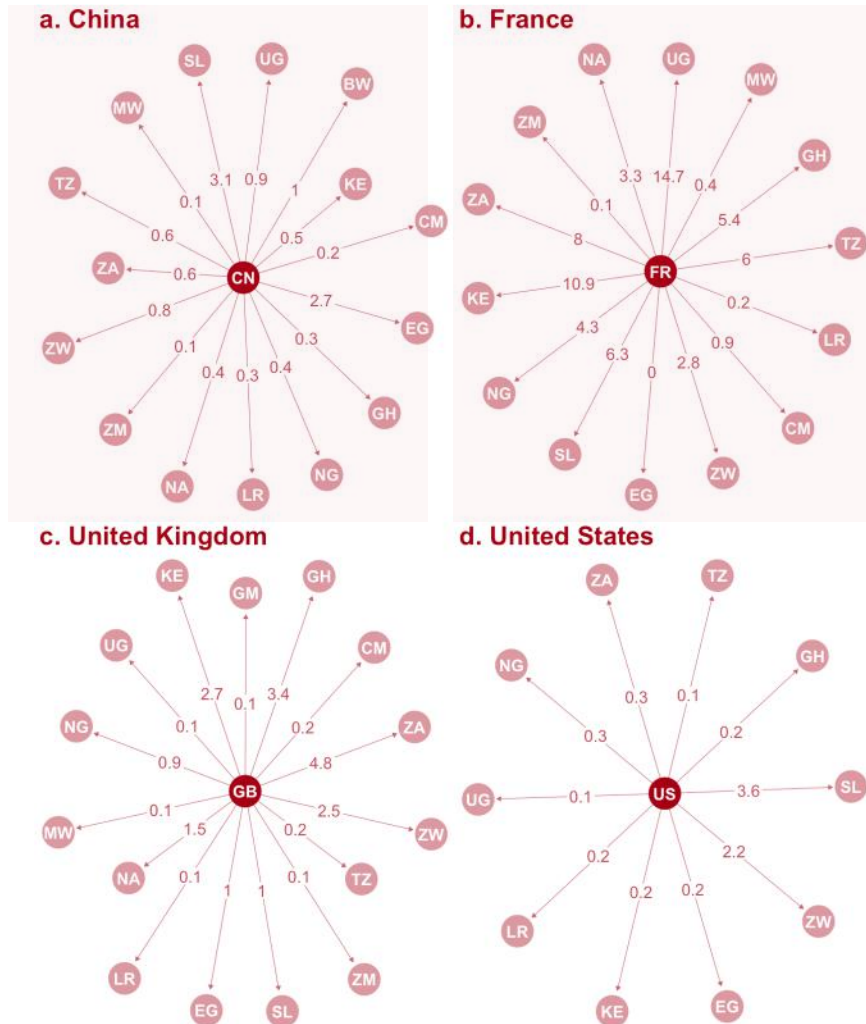
RQ₁: Amount Coverage about China & COVID



Limited news about
China pre-COVID-19;
peaked in April

Global media discuss
China more
XH 225 stories/day
Reuters: 135/day
Average: 1.2/day

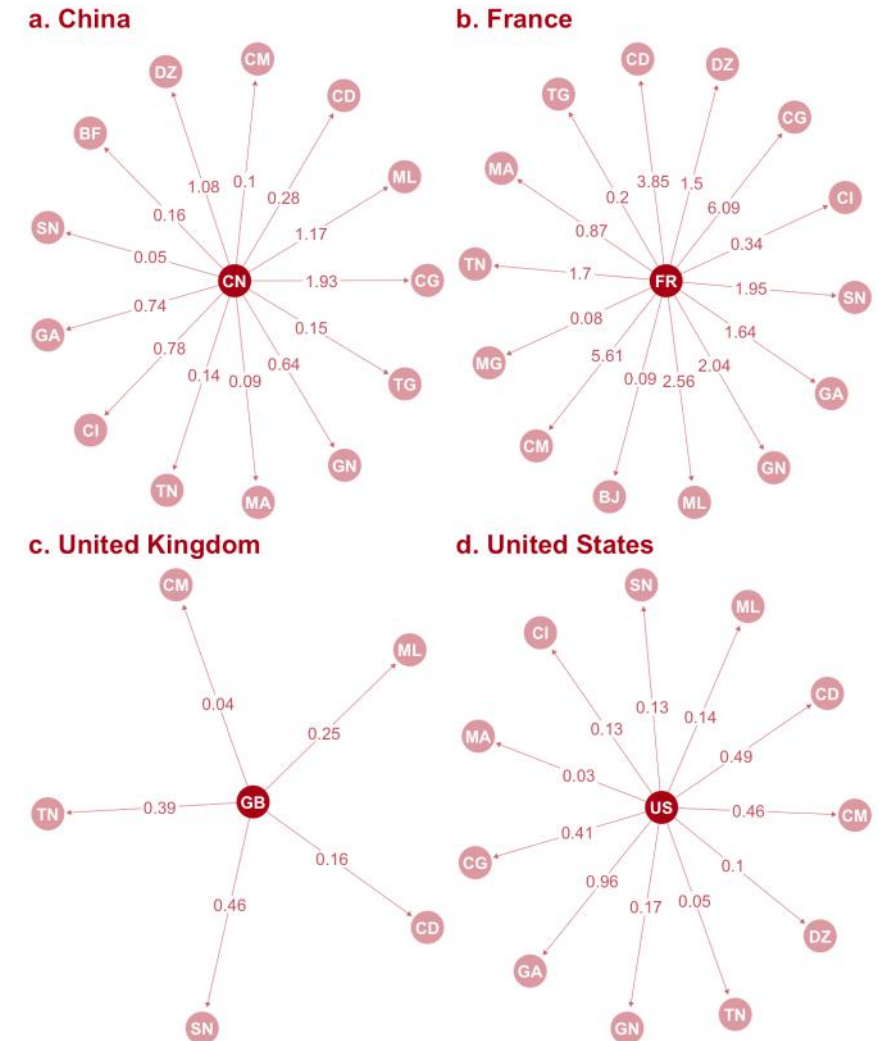
RQ₂ & RQ₃: Quantifying Text Reuse



Around 4% of stories were from foreign sources (AFP, 1st).

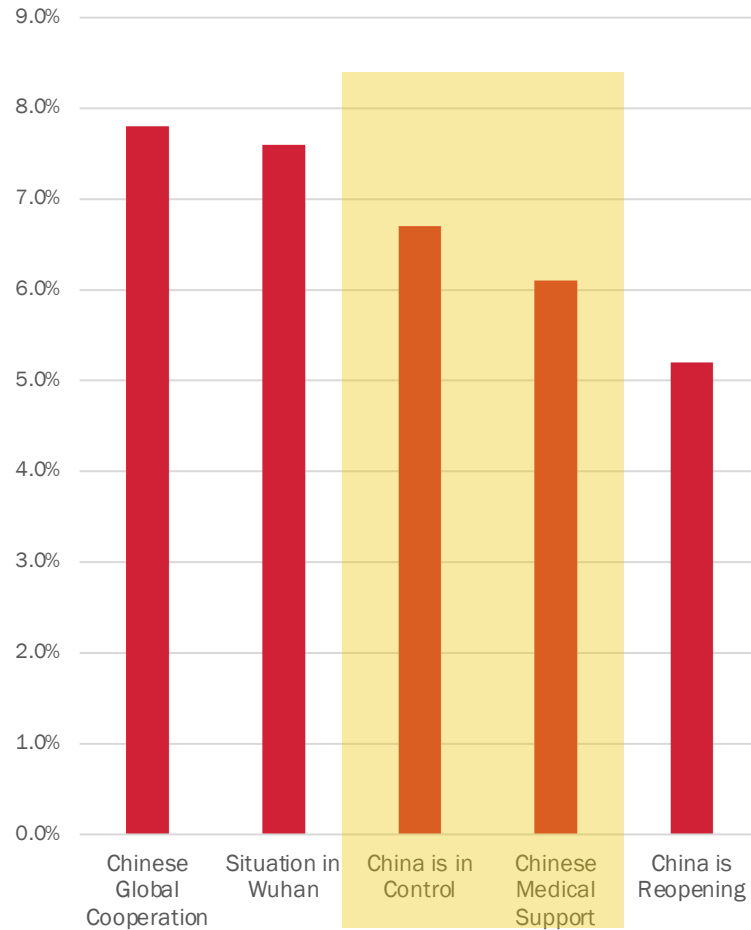
News on China and COVID came from FR (65%), UK (25%) & CN (7%) sources.

More reliance on text reuse on English (<=) that French (=>)

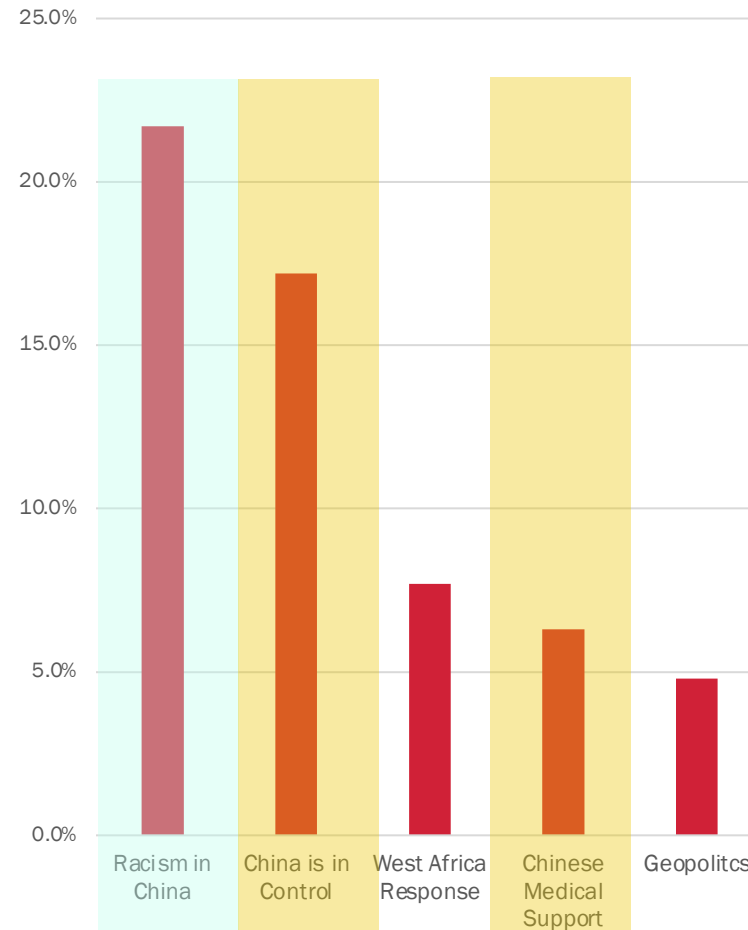


RQ₄: Narrative overlap

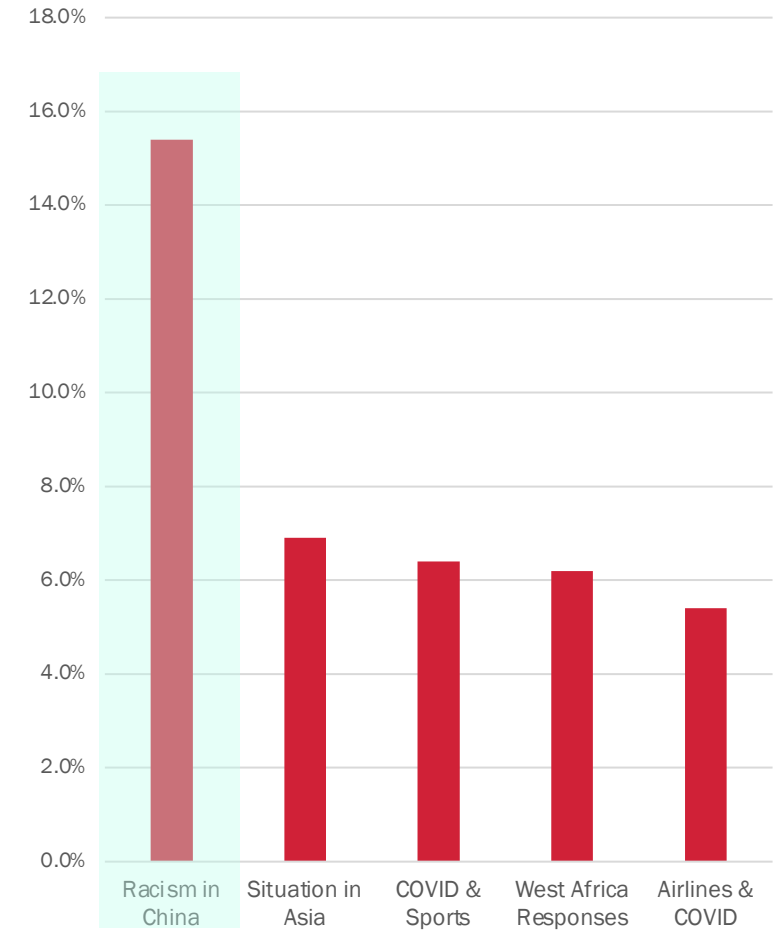
Chinese media



Gambian media



Kenyan Media



Final Ideas I

- This paper found limited evidence of **transnational intermedia agenda setting** :
 - News coverage of China in most African countries tends to be **relatively modest** when no significantly newsworthy “news event” is taking place
 - The use of content published originally by Chinese sources is **very marginal** in both English and French-speaking Africa
 - There is **a gap between the prevailing themes** found in news coverage about China and COVID-19 in Chinese & African media.

Final Ideas II

- Possible explanations:
 - Existing **mistrust among some journalists** in Africa towards Chinese media's content (maybe increased during COVID-19)
 - Preference for **domestic issues** during a pandemic, which might not be covered by Chinese sources.

More information

- (Some of the) data and the code for this study is available in this repository:

https://github.com/danimadrid/africa_china_covid19

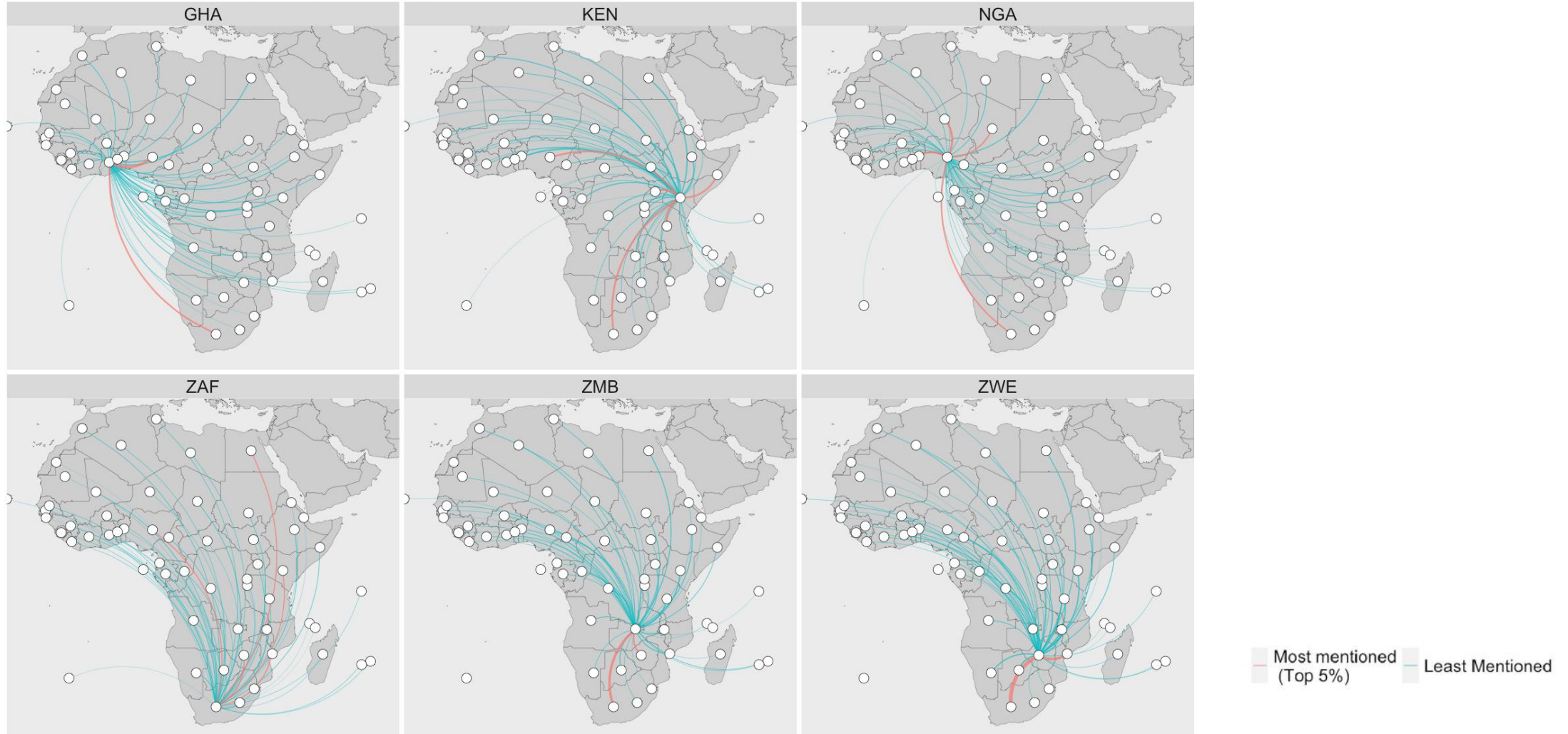
- A paper detailing some of these findings is press:
Madrid-Morales, D. (2021). Who Set the Narrative? Assessing the Influence of Chinese Global Media on News Coverage of COVID-19 in 30 African Countries. *Global Media and China*, 6(2).

<https://doi.org/10.1177/20594364211013714>

Revisiting “Foreign News”

African media covering African news

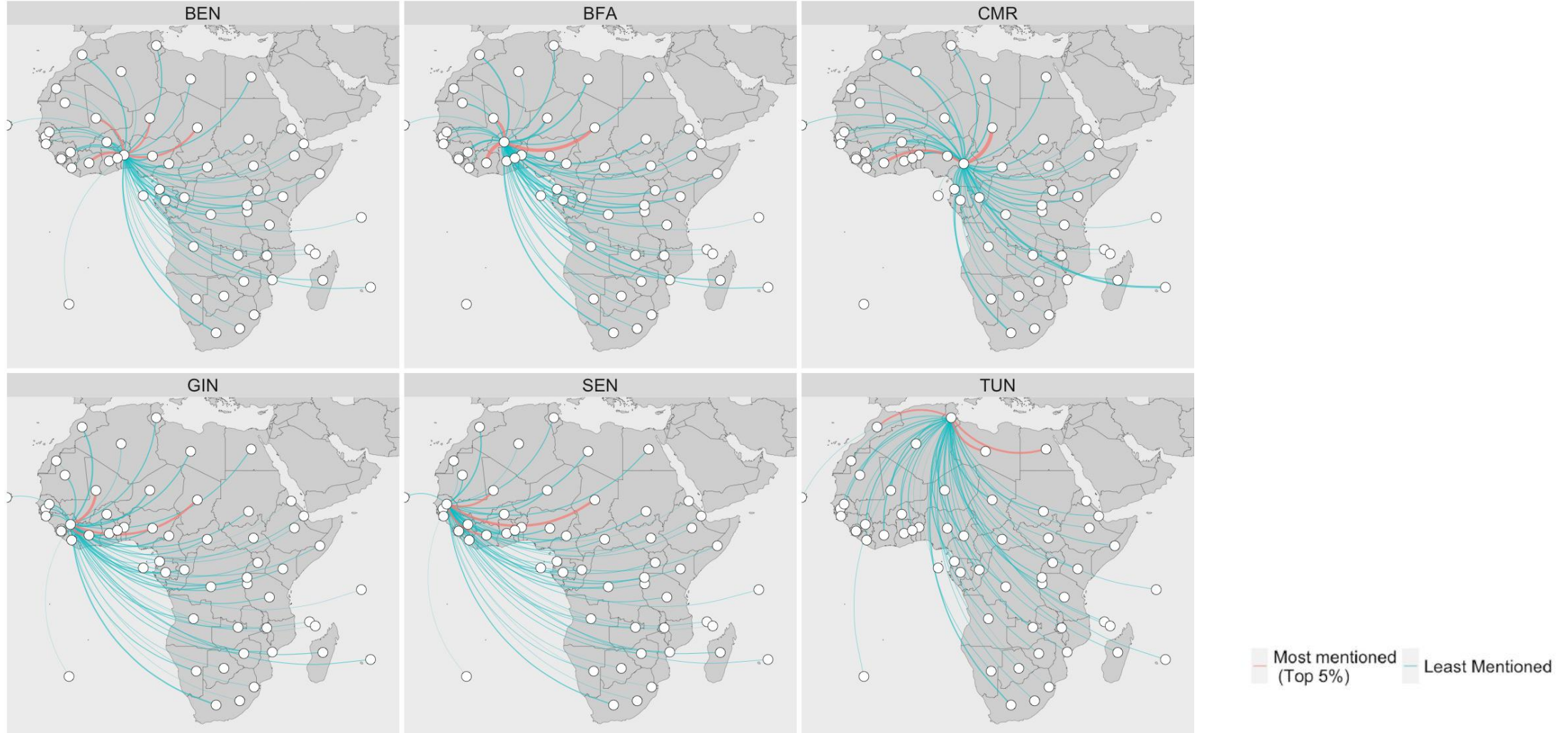
Frequency of country mentions in the media of selected countries (English corpus)
From March 2021 to Mid May 2021 (n = 573,039)



African media covering African news

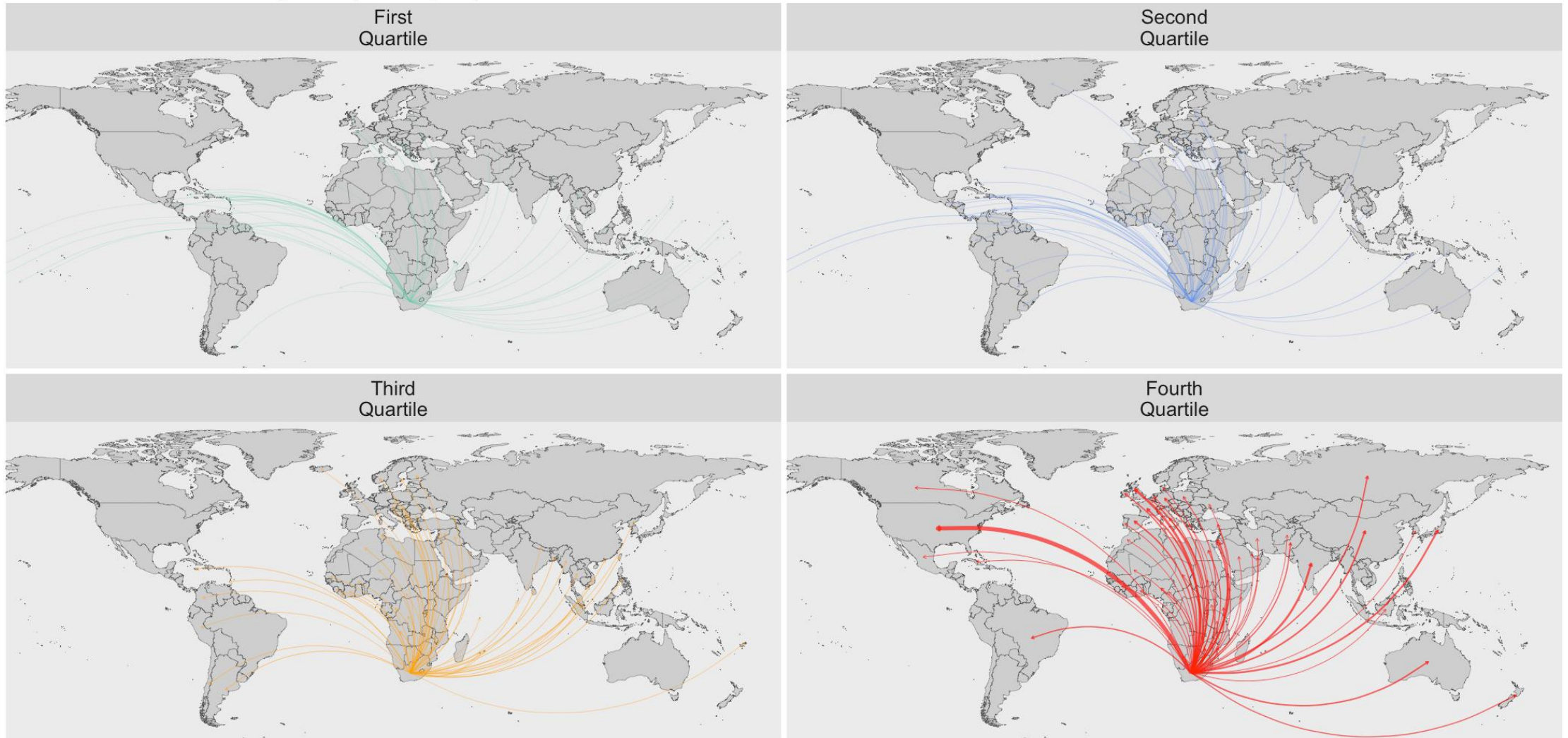
Frequency of country mentions in the media of selected countries (French corpus)

From April 2021 to Mid May 2021 (n = 413,213)



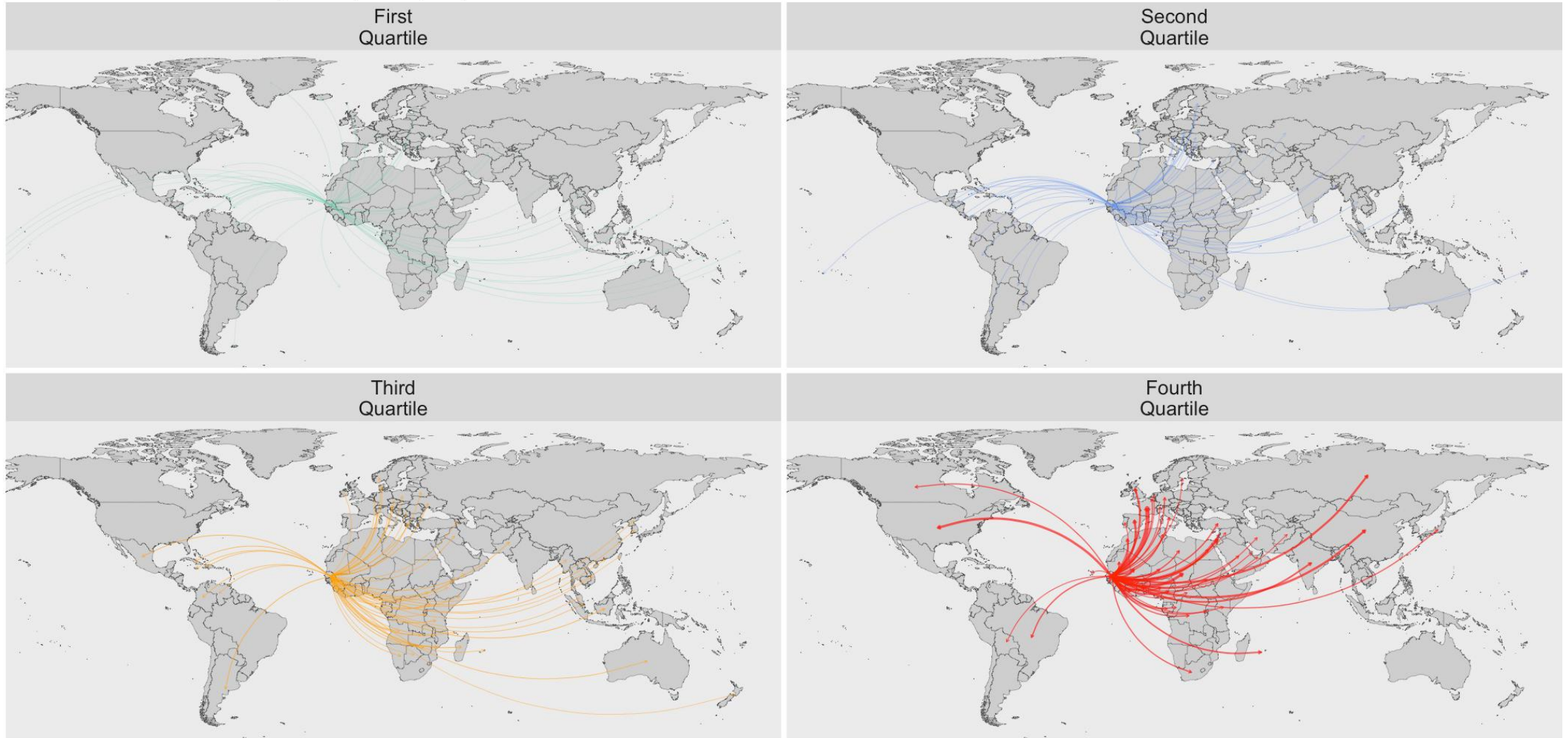
South African media covering the World

From March 2021 to Mid May 2021 (n = 127,141)



Senegalese media covering the World

From March 2021 to Mid May 2021 (n = 88,294)



Next stages for adn

Pilot study with English text [2019]

Beta version of multilingual database [current]

Launch production stage [June 2021]

Opening access to monitor corpus to users [September 2021]

Data processing and insights [pending funding]

Q&A

Old Questions, Big Data: Using the African Digital News (ADN) Corpus to Study News Flows

Dani Madrid-Morales, PhD

dmmorales2@uh.edu

@DMadrid_M

www.danimadrid.net