Tracking Civic Space in Developing Countries with a High-Quality Corpus of Domestic Media and Large Language Models

Supplemental Materials

Table of contents

Appendix A: Data in HQMARC

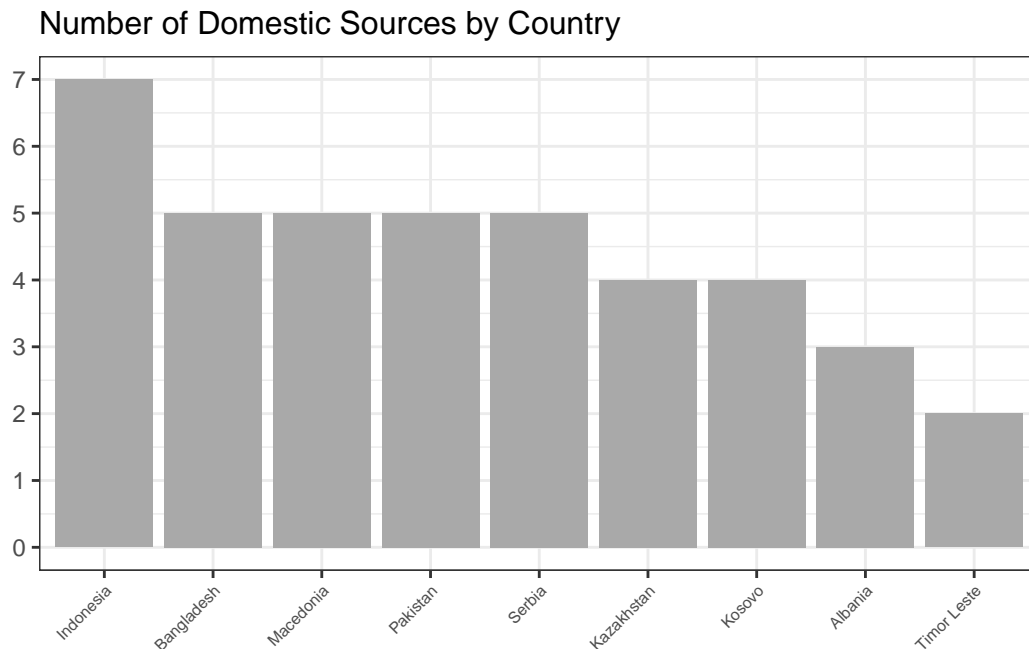## Section 1: Distribution of Domestic Sources in HQMARC

Figure 1: The number of domestic media outlets included in the HQMARC corpus by country.

Section 2: Languages in HQMARC

The *HQMARC* corpus currently includes domestic media outlets publishing in 36 languages. This number will increase as we add new countries and media sources to the corpus. These languages include Afrikaans, Albanian, Amharic, Arabic, Armenian, Azerbaijani, Belarusian, Bengali, Chinese, English, French, Georgian, Hindi, Hungarian, Indonesian, Kazakh, Central Khmer, Kinyarwanda, Kongo, Macedonian, Malay, Nepali, Portuguese, Romanian; Moldavian; Moldovan, Russian, Sinhala; Sinhalese, Spanish; Castilian, Serbian, Swahili, Tagalog, Turkish, Ukrainian, Urdu, Uzbek, Zulu. Below, we provide the languages associated with media outlets in each country using the language ISO codes.

```
**SLV: El Salvador**: es, en

**LBR: Liberia**: en

**KEN: Kenya**: en

**HND: Honduras**: es, en

**KGZ: Kyrgyzstan**: en, ru, kg
```

SEN: Senegal: fr

UZB: Uzbekistan: en, ru, uz

UKR: Ukraine: uk, ru, en

ZMB: Zambia: en

NPL: Nepal: ne, en

CMR: Cameroon: fr

AZE: Azerbaijan: en, az, ru

PAK: Pakistan: ur, en

GEO: Georgia: ka, en

TLS: Timor Leste: tet, pt

TUR: Turkey: tr

ARM: ArmeniaA: hy, ru

AGO: Angola: pt

BEN: Benin: fr

BFA: Burkina Faso: fr

NER: Niger: fr

ETH: Ethiopia: en, am

LKA: Sri Lanka: en, si

BLR: Belarus: be, ru

MRT: Mauritania: fr, ar

TZA: Tanzania: en, sw

UGA: Uganda: en

PRY: Paraguay: es

MOZ: Mozambique: pt, en

SSD: South Sudan: en, ar

ZWE: Zimbabwe: en

RWA: Rwanda: en, fr, rw

NGA: Nigeria: en

GHA: Ghana: en

COD: Congo: fr

MAR: Morocco: ar, fr

TUN: Tunisia: fr, ar

MLI: Mali: fr

GTM: Guatemala: es

XKX: Kosovo: sq

KHM: Cambodia: km, en

MKD: Macedonia: sq, en, mk

KAZ: Kazakhstan: kk, ru, en

SRB: Serbia: sr

MYS: Malaysia: en, ms, zh

PHL: Philippines: en, tl

IND: India: en, hi

MDA: Moldova: ro, ru, en

BGD: Bangladesh: bn

ALB: Albania: sq, en

ZAF: South Africa: af, en, zu

DZA: Algeria: fr, ar

ECU: Ecuador: es

NIC: Nicaragua: es

COL: Colombia: es

NAM: Namibia: en

IDN: Indonesia: en, id

SLB: Solomon Islands: en

HUN: Hungary: hu

JAM: Jamaica: en

PER: Peru: es

DOM: Dominican Republic: es

MWI: Malawi: en

Section 3: Digital News Sources

- International Sources:

aljazeera.com, bbc.com, csmonitor.com, france24.com, nytimes.com, reuters.com,
scmp.com, theguardian.com, themoscowtimes.com, washingtonpost.com, wsj.com, lemonde.fr,
liberation.fr, elpais.com, lefigaro.fr, xinhuanet.com,

Sub-Saharan Africa:

- Africa Regional Sources:

africanews.com theeastafrican.co.ke iwpr.net

- Angola:
opais.co.ao, jornalf8.net, angola24horas.com, portaldeangola.com, angola-online.net
vozdeangola.com, jornaldeangola.ao,

- Benin:
lanouvelletribune.info, news.acotonou.com, lematinal.media, levenementprecis.com,

- Cameroon:
journalducameroun.com, camerounweb.com, 237actu.com, 237online.com, cameroonvoice.com
lebledparle.com, thesunnewspaper.cm,

- DR Congo:
radiookapi.net, lesoftonline.net, acpcongo.com, lephareonline.net, groupelavenir.org
matininfos.net, cas-info.ca, actualite.cd, 7sur7.cd,

- Ethiopia:
addisfortune.news, addisstandard.com, capitalethiopia.com, thereporterethiopia.com, ethiopianm
addisadmassnews.com,

- Ghana:
dailyguidenetwork.com, ghanaweb.com, graphic.com.gh, newsghana.com.gh,

- Kenya:
kbc.co.ke, citizen.digital, nation.africa, theeastafrican.co.ke,

- Liberia:
thenewdawnliberia.com, liberianobserver.com, analystliberiaonline.com
frontpageafricaonline.com, inquirernewspaper.com, thenewsnewspaper.online

- Mali:
maliweb.net, malijet.com, news.abamako.com,

- Malawi:
mwnation.com, nyasatimes.com, times.mw, faceofmalawi.com, malawivoice.com

- Mauritania:
alwiam.info, lecalame.info, journaltahalil.com, alakhbar.info, saharamedias.net

- Mozambique:
correiodabeiraserra.com, canal.co.mz, mmo.co.mz, cartamz.com, verdade.co.mz
clubofmozambique.com, portalmoznews.com, jornaldomingo.co.mz, tvm.co.mz,

- Niger:
actuniger.com, nigerinter.com, lesahel.org, tamtaminfo.com, airinfoagadez.com
nigerexpress.info, journalduniger.com,

- Nigeria:
guardian.ng, thenewsnigeria.com.ng, vanguardngr.com, thenationonlineng.net,

- Rwanda:
newtimes.co.rw, therwandan.com, kigalitoday.com, umuseke.rw,

- Senegal:
xalimasn.com, lesoleil.sn, enqueteplus.com, lasnews.sn, ferloo.com
nouvelobs.com, sudquotidien.sn,

- South Africa:
timeslive.co.za, news24.com, dailysun.co.za, sowetanlive.co.za, isolezwe.co.za
iol.co.za, son.co.za,

- Tanzania:
ippmedia.com, dailynews.co.tz, habarileo.co.tz, thecitizen.co.tz, mtanzania.co.tz
jamhurimedia.co.tz, mzalendo.co.tz,

- Uganda:
monitor.co.ug, observer.ug, newvision.co.ug, nilepost.co.ug, sunrise.ug
eagle.co.ug,

- Zambia:
lusakatimes.com, mwebantu.com, diggers.news, openzambia.com, lusakavoice.com
dailynationzambia.com, zambianewsnetwork.com, zambiaobserver.com,

- Zimbabwe:
thestandard.co.zw, theindependent.co.zw, herald.co.zw, chronicle.co.zw, newsday.co.zw
thezimbabwean.co, zimbabwesituation.com, newzimbabwevision.com, zimlive.com,

Middle East and North Africa

- Morocco:
leconomiste.com, lematin.ma, assabah.ma

- Tunisia:
assarih.com, babnet.net, jomhouria.com, lapresse.tn

- Turkey:
diken.com.tr, t24.com.tr, sozcu.com.tr, posta.com.tr, sabah.com.tr

Eastern Europe

- Eastern Europe Regional Sources:

euronews.com/tag/eastern-europe neweasterneurope.edu balkaninsight.com iwpr.net

- Albania:
gazetatema.net, panorama.com.al, telegraf.al

- Armenia:
azatutyun.am, aravot.am, 168.am, 1in.am, golosarmenii.am

- Azerbaijan:
azeritimes.com, azadliq.info, abzas.org, turan.az, zerkalo.az,
mikroskopmedia.com, xalqcebhesi.az, musavat.com, ru.echo.az

- Belarus:
nashaniva.by, novychas.by, nv-online.info, belgazeta.by, zviazda.by,
sb.by

- Georgia:
ambebi.ge, georgiatoday.ge

- Hungary:
index.hu, 24.hu, 168.hu, hvg.hu, demokrata.hu

- Macedonia:
koha.mk, slobodenpecat.mk, makfax.com.mk, skopjediem.com, novamakedonija.com.mk

- Moldova:
timpul.md, tribuna.md, unimedia.info, voceabasarabiei.md, publika.md,
ipn.md, zdg.md

- Kosovo:
kosova-sot.info, balkaninsight.com, prishtinainsight.com, botasot.info

- Ukraine:
delo.ua, interfax.com.ua, kp.ua, pravda.com.ua, kyivpost.com,
kyivindependent.com

- Serbia:
rs.n1info.com, juznevesti.com, insajder.net, danas.rs, balkaninsight.com

Latin America and the Caribbean:

- Latin America Regional Sources:

elpais.com cnnespanol.cnn.com iwpr.net

- Colombia:
elcolombiano.com, elespectador.com, elheraldo.co, eltiempo.com

- Ecuador:
elcomercio.com, eldiario.ec, elnorte.ec, eluniverso.com, metroecuador.com.ec

- El Salvador:
laprensagrafica.com, elfaro.net, elsalvador.com, diario.elmundo.sv, diarioelsalvador.com,
revistafactum.com, gatoencerrado.news

- Guatemala:
prensalibre.com, republica.gt, lahora.gt, soy502.com

- Honduras:
elheraldo.hn, laprensa.hn, proceso.hn, tiempo.hn

- Jamaica:
jamaica-gleaner.com, jamaicaobserver.com

- Nicaragua:
confidencial.com.ni, laprensani.com, nuevaya.com.ni, articulo66.com, laverdadnica.com,
ondalocalni.com, canal2tv.com, lajornadanet.com

- Paraguay:
abc.com.py, lanacion.com.py, ultimahora.com

- Peru:
elcomercio.pe, gestion.pe, larepublica.pe, ojo-publico.com, idl-reporteros.pe

Asia:

-   Asia Regional Sources:

asiatimes.com asia.nikkei.com iwpr.net

- Bangladesh:
prothomalo.com, bd-pratidin.com, kalerkantho.com, jugantor.com, dailyjanakantha.com

- Cambodia:
kohsantepheapdaily.com.kh, moneaksekar.com, phnompenhpost.com, cambodiadaily.com

- Indonesia:
thejakartapost.com, jawapos.com, kompas.com, mediaindonesia.com, sindonews.com,
beritasatu.com, hariansib.com

- India:
amarujala.com, indianexpress.com, thehindu.com, hindustantimes.com, deccanherald.com,
firstpost.com, indiatimes.com, timesofindia.indiatimes.com

- Kazakhstan:
caravan.kz, diapazon.kz, kaztag.kz, rus.azattyq.org

- Kyrgyzstan:
akipress.com, 24.kg, kloop.kg, super.kg, vb.kg,
kaktus.kg, kaktus.media

- Malaysia:
malaymail.com, nst.com.my, thestar.com.my, utusan.com.my, thesun.my,

```
malaysiakini.com
```

```
- Philippines:
mb.com.ph, manilastandard.net, inquirer.net, manilatimes.net
```

```
- Sri Lanka:
dailymirror.lk, island.lk, divaina.lk, adaderana.lk, lankadeepa.lk
```

```
- Uzbekistan:
fergana.ru, kun.uz, gazeta.uz, podrobno.uz, batafsil.uz,
sof.uz, anhor.uz, asiaterra.info, daryo.uz
```

Section 4: Assessing Outlet Independence

MLP collects extensive news data from various national news sources across countries. Bias is checked for all our sources at a three-month frequency due to:

- **Frequent changes** in coverage of sources.
- **Changes in ownership** of sources.

When we forecast, tracking our accuracy is essential for reporting purposes. Weighting helps enhance the accuracy of MLP civic space forecasting. The bias check exercise assists in assigning weights during the forecasting stage of the MLP pipeline. This is achieved by evaluating the level of independence of all media sources for a country and their respective volumes. If the majority of articles for a country originate from independent sources, no weighting is required.

Section 4.1: Task

- **Assess** and **report** the bias of the media sources we are currently scraping.
- **Formulate** a standardized and formal process to assess bias.

Section 4.2: Bias Categories and Coding Guide

1. Ideology Political ideology of the source is coded as follows:

- 1 = Right
- 2 = Right-centre
- 3 = Centre
- 4 = Left-centre
- 5 = Left

2. Independence vs Pro-regime Determine whether the source operates as a regime mouthpiece, exhibits a bias in favor of political actors, or maintains independence (whether investigative or neutral). It is possible for a source to align with a certain ideology without being influenced by political parties, government, or oligarchs.

- 1 = Pro-regime

- 2 = Opposition
- 3 = Independent

3. Government Owned Indicate if the source is partially or completely owned by the country's government.

- 1 = Government-owned
- 0 = Not government-owned

Process

1. Start by checking the information about a country's newspapers provided in the sources listed in the appendix (e.g., BBC or RSF) to assess and record the bias of each newspaper.
2. If step 1 does not provide enough information, dive deeper into country-specific analyses of the local media, reports on media ownership, etc.
3. If steps 1 and 2 are not sufficient, perform a subjective review of the news reported on the website to get a sense of bias.
4. If steps 1, 2, and 3 are not successful, report as unable to assess independence.

Helpful sources:

- BBC Media Guide country profile: search the web for *Country Name BBC Media Guide*
- RSF- Reporters Without Borders country profile: RSF Country Profiles
- World Newspapers country profile: World Newspapers Country Profiles
- Freedom House: Freedom House Country Profiles

Section 5: Source Weighting

The weighting process adjusts article counts from state-owned and independent sources when aggregating civic space data at the country-month level. In countries where more than 50% of articles come from state-owned sources, the code increases the influence of independent sources by applying a weight based on the ratio of state-owned to independent articles. This ensures that independent media retains influence when state-owned coverage dominates. The weighting is recalculated each month, meaning that the balance of influence can shift over time depending on the mix of articles from state-owned and independent sources in any given month. We impose equal weighting on independent and non-independent sources by multiplying the count of articles for each *RAI* event by a weight calculated by subtracting the number of articles from non-independent sources from the result of dividing the number of non-independent articles by 0.5 and dividing this value by the number of articles from independent sources for each month. Weighting is currently applied to India, Malaysia, Sri Lanka, Angola, DR Congo, Belarus, Philippines, Paraguay, Zambia, and Zimbabwe.

Appendix B: Civic Space Definitions and Keywords

Section 1: Event Definitions

- Activisim

  - Any activity by individuals, community organizations, or NGOs that involve some amount of lobbying, raising awareness, or collective action beyond just the donation of money. Importantly, this does not include protests or activities related to election campaigns.
  - **Example:** "Retired soldiers make up subsidiary of Just Cause.A group of retired soldiers met yesterday morning in a restaurant in the city of San Miguel to learn about the work of the Association Causa Justa, in order to organize a branch in the eastern part of the country.Jaime García, vice president of Causa Justa, explained that this association has been working for a couple of years to defend the interests of retired soldiers, and since then they have organized a branch in San Ana and now they intend to have another group."

- Arrests

  - An institution within the government-controlled security apparatus –i.e. the police, the military, or other– apprehends people or groups of people who are part of an opposition movement or party, a civil society organization, a media organization, or a protest.
  - **Example:** "The Police in Abia on Wednesday arrested 51 persons suspected to be members of the proscribed Indigenous People of Biafra, claiming they 'operate as members of Judaism,' in Umuahia."

- Censor

  - The government actively prevents free speech by individuals in the media, in public or online. This includes banning certain content from individual speech or news stories, dictating how certain concepts or people can be referred to in public speech, or directly dictating agenda setting for media organizations. This category also includes the government censoring internet websites, internet shutdowns, fines on independent media, limitations on foreign ownership of media outlets, and political actors gaining influence within media organizations. Magnitude is a scale (see below; try to determine the importance of the target with information given).
  - **Example:** "The Tanzanian government has suspended newspaper The Citizen for seven days after the publication ran a story on the falling value of the Tanzanian shilling."

- Cooperation

  - Domestic political, social, or business actors collaborate on one or a range of issues or demonstrate an intent to do so. This does not include intragovernmental cooperation, except for local and federal cooperation. Cooperation indicates a willingness for domestic actors to work together to resolve important issues. Cooperation involving international actors is not a civic space event, but is an RAI event.
  - **Example:** "Opposition movement NJPE has pledged to work together with President Muhammad to reduce rampant poverty in the capital's outskirts."

- Corruption

  - The abuse of entrusted power for private gain. This includes street-level corruption (for instance, bribes to police), high-level corruption (in public contracting, for instance), and legal actions around corruptions.
  - **Example:** "Ex-President Sarkozy gets Jail Sentence for Corruption in France"

- Coup

  - Changes in government or persistence of government that are not in the rules of succession or transition. Coups, refusing to cede power, or a power grab after an unfair election are examples of this event type. Peaceful government transitions are nonviolent transfers of power or legitimate continuity of government elected by democratic means and accepted by a majority of political forces. Code peaceful change as '1' in the direction column. Speech is the pledge to accept the result of an election (a threat of a coup falls under 'political threat').
  - **Example:** "Forces loyal to Turkey's president quashed a coup attempt in a night of explosions, air battles and gunfire that left at least 161 people dead and 1,440 wounded, yesterday. President Recep Tayyip Erdogan vowed that those responsible "will pay a heavy price for their treason".

- Defamation Case

  - Cases in which an individual or a group related to an opposition movement or party, current or former government officials, a civil society organization, a media organization, a business leader, a member of a minority group, or a protest are accused, usually by a government official, of: directly defaming/libeling/slandering the government or one or some of its members. Alternatively, these are: strategic lawsuits against civic actors to intimidate public participation. This category is a subset of the 'Legal Action' event type (see below).
  - **Example:** "A Phnom Penh court on Friday found veteran opposition chief Sam Rainsy guilty of defaming Prime Minister Hun Sen and ordered him to pay damages of $1 million, the latest blow to an opposition crippled by legal cases this year."

- Disaster

  - This category includes: (1) natural disasters such as earthquakes, hurricanes, floods, famines or food crises or any catastrophic event that results or may result in serious damage. loss of life, and/or political implications; (2) infrastructural accidents that can endanger the population –a dam or pipe bursting, the sinking of a ferry, etc.) The goal is to focus on the types of events that might reflect on government accountability and civic space, not minor events like traffic accidents. Note: this includes deaths or economic strife caused by disasters, including reports on deaths/infections caused by COVID-19.
  - **Example:** ROMBO district residents in Kilimanjaro region are in danger of facing food shortage if concerted efforts to control wildlife in consuming crops are not employed.

- Election Activity

- Reporting on regular electoral activities including the introduction of candidacies, the conduct of campaigns, the announcement of results, and the formal transfer of power for public office. This includes situations where the incumbent party/candidate retains power. Normal election results are an example of this event type. Specifically, this includes peaceful government transitions that are nonviolent transfers of power or legitimate continuity of government but does not include coups, which are accounted for separately.
  - **Example:** "The President of the Executive Board of the Serbian Progressive Party, Darko Glisic, said that the SNS won a convincing victory with 60.2 percent of the votes, based on the processed 221 out of 231 polling stations where the voting for the parliamentary elections in Serbia was repeated."

- Election Irregularities

  - The altering or attempted altering through legal or extra-legal manipulation of the conduct, rules or results of elections or electoral processes. This could include rescheduling/postponing/cancelling regularly-scheduled elections, calling irregular elections or constitutional referenda, hampering the work of independent election observers, eroding the autonomy/authority of the electoral commission, as well as any other institutional change that directly affects the electoral process. This additionally could include the cancellation of party lists and the boycott of elections by opposition, as well as calls for recounts or other forms of contesting elections. Finally, this includes vote-buying, intimidation, or vote-rigging. This does not include standard election proceedings.
  - **Example:** "Nigeria's electoral authority has delayed presidential and national assembly elections by one week amidst protests from the two main opposition parties. The government alleged logistical problems with ballot delivery to justify the delay."

- Legal Action

  - Legal action refers to the prosecution or investigation of criminal activity or disputes over government authority, as well as the trials, convictions or dismissals that result. This event is related to civic space if the targeted people are part of an opposition movement or party, current and former government officials, a civil society organization, an NGO, a media organization or member, a (legal or illegal) business leader, member of minority group, or a protest. This specifically does not include arrests, which are defined as its own event type. Defamation cases are a subset of 'Legal Action' (see above).
  - **Example:** "Kosovo President is Indicted for War Crimes for Role in War with Serbia."

- Legal Change

  - Legal change refers to any changes or proposed changes in the laws of a nation in such a way as to affect civic space. This includes legal restrictions on speech, political groups, NGOs, and the changing of constitutions as well as changing of the powers of the government. This also includes some restrictions on assembly, but does not include curfews and other martial law declarations, which are specifically covered by other event categories.
  - **Example:** "Venezuela's constituent assembly yesterday unanimously passed a law that mandates punishment including a prison sentence of up to 20 years for anyone who instigates hate or violence on the radio, television or via social media."

- Cooperation

- Domestic political, social, or business actors collaborate on one or a range of issues or demonstrate an intent to do so. This does not include intragovernmental cooperation, except for local and federal cooperation. Cooperation indicates a willingness for domestic actors to work together to resolve important issues. Cooperation involving international actors is not a civic space event, but is an RAI event.
  - **Example:**

- Lethal Violence

  - Any action of aggression by a government entity, organized group or individual that results in the death of one or more people, excluding crimes of passion.
  - **Example:** "The Kaduna state government, Friday evening, disclosed that 33 women and children were killed by rebels in Kajuru local government of Kaduna state, less than twenty-four hours to the conduct of the presidential and parliamentary elections."

- Martial Law

  - The executive branch declares a state of emergency or suspends, temporarily or indefinitely, the ability of citizens to gather or protest against the order.
  - **Example:** "On Wednesday, President Duterte approved the extension of martial law in the country's volatile south by a year due to continuing threats by Islamic State group-linked militants and communist insurgents."

- Mobilize Security Forces

  - An event in which the government mobilizes police forces, military troops or government-affiliated militias.
  - **Example:** "More than 500 security personnel have being mobilised for Saturday's governorship election in Sokoto State, says the state commissioner of police (CP), Alhaji Adisa Bolanta."

- Non-Lethal Violence

  - Any action of aggression by a government entity, organized group or individual that physically harms one or more people or property but does NOT result in death, excluding crimes of passion.
  - **Example:** "At least four persons were injured at Oruk Anam Local Government Area of Akwa Ibom State on during yesterday's national election. Reports claim that unknown assailants attempted to snatch ballot boxes while voting was still ongoing."

- Protest

  - Planned or spontaneous public mobilization of a large group of people. Labor strikes, political rallies and riots are also included in this category. When protest events involve deaths, they are coded as lethal violence, but this does not apply for non-lethal violence incidents.
  - **Example:** "A reported two thousand people took to the streets yesterday in Nairobi to protest rising fuel prices, which have doubled since the beginning of the year."

- Purge/Replace

– Purge/replace refers to the removal, firing or resignation of individuals from a government position or the replacement of previously removed individuals, including when the removals/firings themselves are not directly mentioned in the article. This includes the resignation of the chief executive. This description applies to purging of targets such as the bureaucracy, courts, military, police, state-owned companies, or members of political parties, among others.
– **Example:** "Poland's government carried out a sweeping purge of the Supreme Court on Tuesday night, eroding the judiciary's independence, escalating a confrontation with the European Union over the rule of law and further dividing this nation."

- Raid

  – Individuals or organizations are assaulted or aggressively coerced. Their property may be encroached or damaged as a result. Examples include a raid on newspaper offices. Victims themselves suffer no physical harm. This category also includes the government shutting down opposition organizations, NGOs, etc.
  – **Example:** "A Vanguard newspaper office located on Bassey Duke Street was, yesterday afternoon, raided by hoodlums, who carted away large sums of money and destroyed computers and other equipment."

- Threats

  – A statement of a clear and explicit intention to inflict pain, injury, damage, or other hostile action on an individual or organization. Our targets and actors of interest are part of an opposition movement, the media, a political party, government, or a civil society organization.
  – **Example 1:** "A top Kenyan newspaper published a fake death notice of a prominent opposition financier on Wednesday, a bizarre error that rights groups interpreted as another sign of an anti-democratic slide. The Daily Nation apologized by mid-morning for publishing the funeral announcement for businessman Jimi Wanjigi, whose picture, history and family details were used but whose name was slightly altered. The paper said the ad was published in error and it was working with police to uncover who placed it. After a week of arrests of opposition politicians and a crackdown on independent media, a prominent rights campaigner said the announcement amounted to a death threat to Wanjigi, who funded opposition leader Raila Odinga's election campaign last year and whose house was raided by police in October."
  – **Example 2:** "Nigeria's main labor unions threatened a nationwide strike over recent increases in gas prices."

Section 2 Keywords and Their Use

This document lists the categories where keywords are currently being deployed as well as a brief explanation for their use.

Legal Action

- Keyword lists

  – Keyword list one: case | lawsuit |sue | suit | trial | court | charge | rule | sentence | judge

– Keyword list two: defamation | defame |libel | slander | insult | disparage | lese majeste | lese-majeste | lese majesty | reputation

- Purpose of Keywords

  – The purpose of these keywords is to move articles from legal action that are actually defamation case to the appropriate category. There is no longer a defamation case category, it is now entirely a subset of legal action. The first set of keywords filter out any instances where there are accusations of defamation/libel/slander that are not actual cases, but merely statements. The second set of keywords ensures that the cases are actually related to defamation, since we found that oftentimes, non-defamation cases where being assigned to this event category. The process is two-fold, requiring a key word from both lists.
  – If keyword from both lists are not present, the article is left as legal action

Censor

- Keyword list: Freedom | assembly | association | movement | independent | independence | succession | demonstrate | demonstration | repression | repressive | crackdown | draconian | intimidate | censoring | controversial | censor | muzzle | restrictive | restrict | authoritarian | non-governmental organizations | NGOs | media | parties | civil society | opposition | critics | opponents | human rights groups | arbitrary | stifling | ban | strict | boycott | protests | dissent | demonstrators | journal.* | newspaper | media | outlet | censor | reporter | broadcast.* | correspondent | press | magazine | paper | black out | blacklist | suppress | speaking | false news | fake news | radio | commentator | blogger | opposition voice | voice of the opposition | speech | broadcast | publish | limit.* | independ.* | repress.* | journalist | newspaper | reporter | internet | telecommunications | magazine | shut down | broadcast | radio
- Purpose of Keywords

  – The purpose of these keywords is to filter out instances where restrictions are applied that are not censorship. This was created primarily in response to closings of schools, businesses, and government offices in relation to COVID-19, which were frequently classified as censorship.
  – If one of the keywords is not present in the article, it is reclassified as -999

Legal Action/Purge/Arrest

- Keyword list

  – embezzle | embezzled | embezzling | embezzlement | bribe | bribes | bribed | bribing | gift | gifts | gifted | fraud | fraudulent | corrupt | corruption | procure | procured | procurement | budget | assets | irregularities | graft | enrich | enriched | enrichment | laundering | fraudulent.

- Purpose of Keywords

  – The purpose of these keywords is to assign a second event category to those articles in arrest, purge, and legal action that also feature corruption. Articles in those categories that feature these words are double counted as both corruption and the original category.

Corruption

- Keyword lists

  - For arrest: arrest; detain; apprehend; capture; custody; imprison; jail
  - For legal action: legal process; case*; investigate*; appeal; charged; prosecute*; case; lawsuit; sue; suit; trial; court; charge; rule; sentence; judge
  - For purge: resign*; fire*; dismiss; sack; replace; quit.

- Purpose of Keywords

  - The purpose of these keywords is to assign a second event category to those articles in corruption that also feature arrests, legal action, and purges. Articles in those categories that feature these words are double counted as both corruption and the original category.