

A Text Analysis-Based Study of Climate Change News Across Countries

I. Introduction

Climate change is a global issue that affects everyone in all countries, and it requires urgent international commitment and action. However, countries have different perspectives and priorities, and reaching an international consensus and agreement on action remains a challenge. Understanding the different attitudes and agendas regarding climate change that each country has would inform negotiations, the policy making process, and international agreements.

We will analyze the public attitude toward ‘climate change’ in countries around the world and try to find broad patterns and trends. We will do this specifically by analyzing how climate change is portrayed and the contexts it is embedded in in newspaper articles from several countries around the world. Newspapers are a more objective media outlet (as opposed to blogs, etc.) that is also reflective of both current public consciousness (as opposed to lengthier books) and government priorities (as opposed to policy papers, etc.). For the purposes of the study, countries are actors operating in the climate change arena, and the media of each country reflects its public attitude.

Due to the large number of articles in the sample, we will utilize methods of natural language processing to conduct the analysis. Computer algorithms will find patterns, similarities, and differences in the large amount of text data that we cannot read, let alone analyze. We will then imbue machine findings with meaning through human analysis.

II. Methodology

A. Natural Language Processing

We will mainly use three methods of natural language processing to analyze the text data, namely, word embeddings, topic modeling, and semantic analysis. Word embeddings will show similarities and differences at the word-level, which we will extend to the document level. Then through topic modeling, we will discover some of the broad themes that are present throughout the corpus. Finally, semantic analysis will give us a sense of the most common actors and some of the actions and objects they take on in the climate change space.¹

B. Sampling

In order to compare attitudes toward climate change across the globe, we chose to gather news items that maximized the geographical spread through locations, while also having enough from a specific place to be able to find the location’s distinct voice in aggregation. As a means to be able to aggregate evenly through regions, we decided to exclusively use news in English. While this may limit sample availability from countries whose main or common language is not English, we found at least

¹ The Python package nltk was used for processing the text data. The package gensim was used for word embeddings and topic modeling, while the Stanford Core NLP was used for semantic analysis.

one country in each region to have a healthy corpus of English language newspapers.² Further, terms like ‘climate change’ may cover different conceptual spaces in different languages, and unifying analysis through just one language allows us to circumvent this problem as much as possible.

The countries we chose to gather news from are: the United States, the United Kingdom, Nigeria, Kenya, Australia, China, India, and the Middle East.³ We limited the time frame to the year 2017, in order to get the most current perspectives on climate change. We used the global news database NewsBank, and built a custom web scraper that simulated a user’s exploration through the website to respect licensing constraints. Since climate change is the topic of this study, we gathered newspaper articles from across all available publications from each country that was related to the keywords ‘climate change’. Taking these constraints in mind, we gathered about 7000 newspaper articles to be used in our sample, comprising of about 1000 articles for each of our countries. Since Nigeria and Kenya had significantly less articles to sample from, we sampled 500 articles from each.

We found the keyword search on ‘climate change’ to result in a good sample corpus. Most articles had the collocation ‘climate change’, while a minority of them covered environmental topics without specific mention of the terms.⁴ There were a lot of duplicate articles published in different newspapers (especially in small, local publications). We deleted almost 1000 duplicates, and pared down our sample corpus to about 6000. Table 1 shows the breakdown of articles by country.

Country	Articles
Australia	684
China	837
India	946
Kenya	449
Middle East	985
Nigeria	492
UK	654
US	889

Table 1: Number of articles per country

III. Word Embeddings

A. Document Similarity

Different societies and groups of people will produce different artifacts based on their values, priorities, agendas, and circumstances. A first approach towards understanding what these differences between different productions of news items regarding climate change are is to compare the end product itself across different societies. Especially given the ever-more interconnected nature of our world, a question worth asking is which countries produce newspaper articles that are more similar to

² Unfortunately, Latin America did not have a thriving English language newspaper scene and was not included in this study.

³ Due to lower English news availability in separate countries in the Middle East, we decided to represent an aggregate of countries in the region as a single country. Thus, our “Middle East” region encompasses Egypt, the United Arab Emirates and Saudi Arabia. This mix of countries allows for both coherence as Arab nations and divergence as different states, a dynamic probably present in some other countries.

⁴ A small number had the words ‘climate’ and ‘change’ separately in an entirely different context, but the number is probably too small to impact analysis.

one another. This will provide the most general sense of similarities and differences between countries, and ground later analyses.

Word embeddings is a technique that maps words in a text into a multi-dimensional space. While there are multiple approaches to creating this space, the objective is for the distance between words in this space to reflect the distance in meaning between the mapped words. Word2Vec is a machine learning algorithm that takes a corpus of text and creates a model of this mapping. Building on this, Le & Mikolov (Le 2014) created a model in which the documents themselves are mapped to a multi-dimensional space. This document-level modeling allows for comparisons between the documents themselves, and not just individual words within the documents. Further, we can aggregate the documents by country and compare similarities between the countries. Figure 1 shows the result of this comparison, using a cosine similarity measure between the vectors that represent each country.

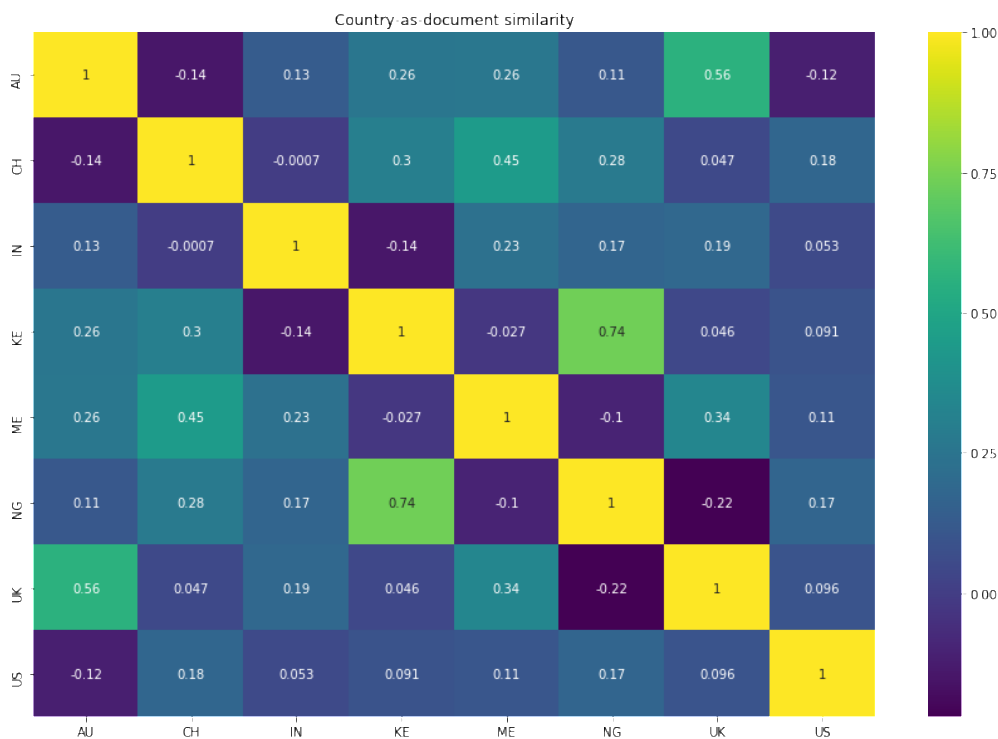


Figure 1: Document-level similarities between countries

The similarity scores exhibit a complex interaction. Nigeria and Kenya, two developing African countries, possess the highest similarity (0.74) among all the pairs of countries. Furthermore, The UK and Australia, two developed “Western” countries, also show a high level of similarity (0.56). However, wealth and culture do not seem to determine similarities among countries. China is surprisingly quite similar to the Nigeria, Kenya, and the Middle East (scores between .28 ~ .45). On the other hand, the US and India each seem to be singing its own tune; neither are very similar to any of the other countries. The Middle East may have the least distinctive voice, as it is moderately similar to all countries.

The previously described approach, however, does not take into account the variability of content within each country. Climate change-related news items produced by a country are not expected to reflect strictly one point of view or one area of interest. To examine the relationship between the

production of news items between countries and within the countries themselves, while taking into account this variability, we compare every document in a country to every other document in each of the countries, including the country itself, and calculate the mean similarity. A country's mean document-document similarity score gives us an indicator of whether its news is generally uniform or diverse. For two different countries, the measure would point to whether the news articles are generally similar. A byproduct of using documents as vectors that represent the meaning behind the article is that if the cosine difference is zero, it would point to the articles being orthogonal, just different, in other words, whereas a positive or negative score indicates that the meaning is either similar or contrary.

Figure 2 shows the result of applying the method described above, comparing every pair of documents in the corpus and obtaining the mean by country. Only the means statistically significantly different from 0 are shown. Foremost, we can see that documents are actually quite different from one another, since all the scores hover around 0. This is not surprising, since articles use different words and are of different lengths. Looking deeper into each country's mean, we can see that Nigeria presents the most uniform voice out of all the countries. We can also see that US news has a negative mean score to news in many countries, including India, Kenya, Nigeria and Middle Eastern countries, who all present positive mean similarities with one another. News in the United States seem to have meaning that is opposed to news in these countries.

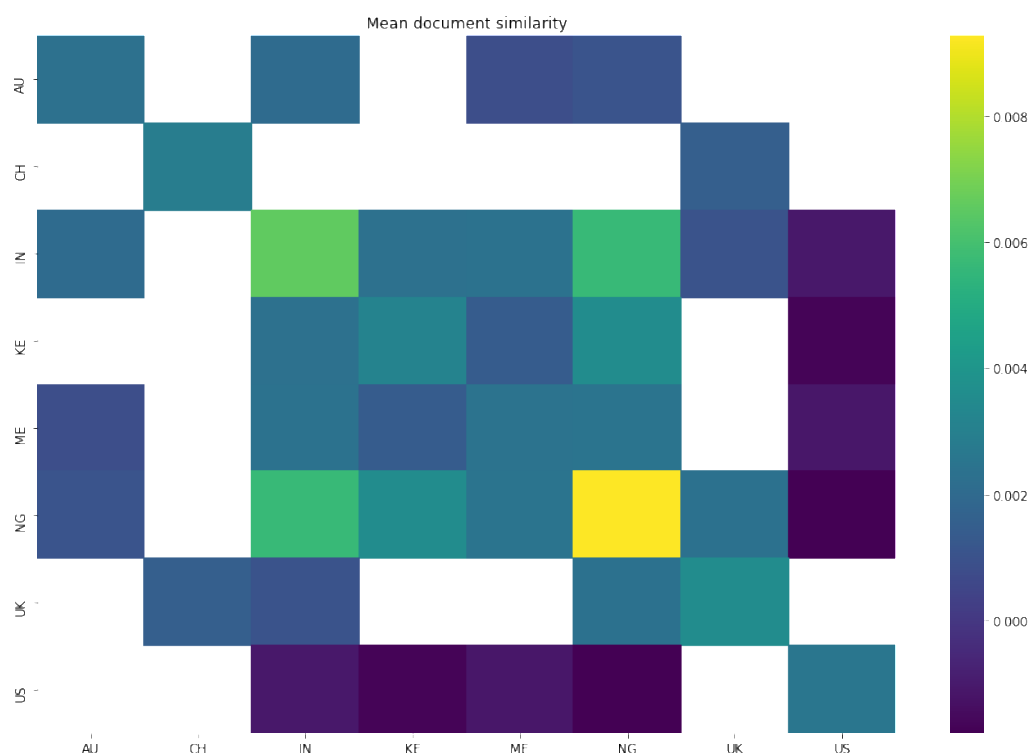


Figure 2: Mean document-to-document similarities

B. Dimensionality

The relationship between words that is encoded in the distances in the word embeddings are also useful in discovering relationships encoded in the text corpus. For example, word embeddings have been used to uncover patterns of gender bias present in text (Caliskan 2017). We use the same method to answer questions regarding the relationships of the countries themselves in the aggregate of all the news items we sampled.

Guided by our initial exploratory search, we decided to explore how the countries are associated to concepts along three different dimensions (shown in Figure 3):

- Outlook: A semantic dimension encompassing hopefulness to hopelessness. The words to test association to are “solution”, “hope”, “action”, and “goal” on one end, with “risk”, “fear”, “disaster”, and “failure” on the other end.
- Veracity: Given the presence of some news regarding climate change denialism, this is a semantic dimension in which we try to see which countries are more associated with scientific consensus versus doubt. The associated words in each end are <“research”, “report”, “science”, “scientist”> and <“hoax”, “alternative”, “doubt”, “skeptic”>, respectively.
- Leadership: This dimension maps the relationship of countries ranging from closeness to “leader” or to “supporter”.

On the outlook dimension, China is associated with positivity and Australia and the UK with negativity. This could be an artifact of a number of things. As we will see in the semantics section, the majority of the mentions of a country’s name (except for the US) comes from the news articles produced by the country itself. China’s status as an outlier in this dimension is perhaps due to state influence,⁵ where the government mandates positive portrayals of China in a leadership role in the press. Countries like Nigeria and Kenya seem to be more fearful or pessimistic concerning the climate change outlook.

Graphing the countries along the veracity spectrum portrays America as alone in climate change denialism, whereas the rest of the countries are more aligned towards scientific consensus. The countries most associated with the positive end of this spectrum are Kenya and Nigeria, which have the lowest GNI per capita in our sample (see Table 12). As we will see in a following section, these countries include more mentions of climate change causing catastrophes, perhaps influencing this belief through tangible effects.

Mapping countries in alignment in a spectrum from ‘leader’ to ‘supporter’, it is clear that news items relate China and America to leaders, whereas the UK is related to more of a supportive role. Keeping in mind that these are all articles that mention climate change, it is interesting to note the rise of China in this arena, even if taking into account the caveats regarding press freedom expressed above.

⁵ Freedom of the Press Index marks China as a country in a “Very Difficult Situation” for reporters (Reporters Without Borders 2018).

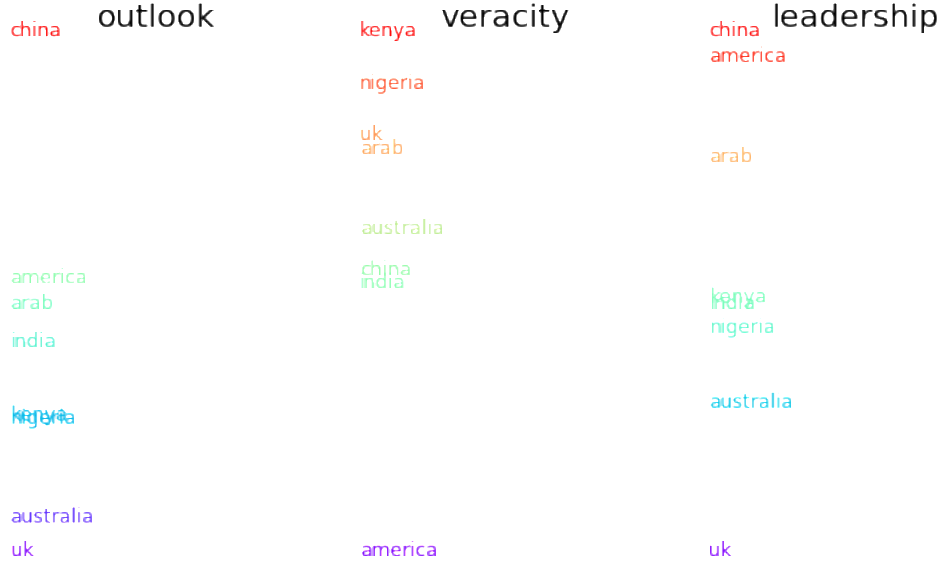


Figure 3: Mapping word embedding-related distances along semantic dimensions

IV. Topics Within Climate Change

While the corpus consists of a collection of articles loosely related to climate change, we know that climate change affects many aspects of society. Topic modeling can help to discover broad themes that run through this loosely related collection of documents, and further reveal the foci of each country’s climate change agenda. We will use a Latent Dirichlet Allocation (LDA) model to determine the topics within the corpus. The LDA model works under the intuition that a hidden generative process produced the collection of documents. It assumes that the corpus consists of a set of topics, and that each document has a different distribution for each of these topics, and finally that these distributions produced each of the words in each of the documents. The LDA algorithm uses the observed words in the documents to infer the most probable hidden topic structure, including the distribution of words within a topic and the distribution of topics within each document (Blei 2012).

We tried many parameter specifications of LDA topic modeling and decided that three topics provided the most coherent and consistent topics.⁶ Table 2 shows the distribution of the most frequent twenty words for each of the three topics, which we labeled ‘politics’, ‘environment’, and ‘development’. Some of the top words in the ‘politics’ topic were such words as ‘Trump’, ‘state’, ‘US’, ‘agreement’, ‘accord’, and ‘plan’, showing the political nature of climate change and the focus on agreed and coordinated action needed by states the international community (agreement and accord most likely refer to the 2015 Paris Agreement). The second topic, corresponding to the ‘environmental’ sphere, features words like ‘scientist’, ‘level’, ‘rise’, ‘weather’, ‘research’, ‘impact’, and ‘flood’, and

⁶ The topics and their coherence levels fluctuated greatly at different values of number of topics, max data-frame (threshold for ignoring words that appear too often), max features (number of words to use in modeling), normalization type (stemming and/or lemmatization of words), alpha (level of mixture of topics within one document), and beta (level of overlap of words in different topics). However, the topics converged to a high level of consistency when certain common but non-informative words were taken out of the corpus, in addition to the basic English stop words in the nltk python package. These words were: ‘climat[e]’, ‘chang[e]’, ‘like’, ‘year’, ‘mr’, ‘say’, and ‘could’. We then chose the number of topics that produced the most coherent topics.

captures the effects climate change has on the natural environment and some of the scientific efforts made in this direction. The third topic covers the economic ‘development’ aspects of climate change. Top words include ‘develop’, ‘china’, ‘food’, ‘nigeria’, ‘water’, in addition to words like ‘work’, ‘make’, ‘project’, and ‘build’, reflecting how climate change features in today’s development agenda.

Politics		Environment		Development	
<i>Word</i>	<i>Probability</i>	<i>Word</i>	<i>Probability</i>	<i>Word</i>	<i>Probability</i>
trump	0.0248	time	0.0094	develop	0.0228
state	0.0192	make	0.0086	china	0.0156
us	0.0171	scientist	0.0082	nation	0.0133
nation	0.0128	rise	0.0082	food	0.0105
agreement	0.0123	human	0.0081	nigeria	0.0095
report	0.0101	us	0.0080	need	0.0094
accord	0.0090	level	0.0077	water	0.0094
plan	0.0088	weather	0.0077	region	0.0093
develop	0.0085	state	0.0075	work	0.0087
need	0.0083	nation	0.0075	use	0.0083
action	0.0082	per	0.0074	research	0.0079
carbon	0.0081	need	0.0071	africa	0.0079
environment	0.0077	impact	0.0071	state	0.0075
support	0.0075	flood	0.0070	area	0.0074
water	0.0074	use	0.0069	make	0.0073
power	0.0067	research	0.0069	project	0.0069
make	0.0065	carbon	0.0067	time	0.0069
work	0.0064	even	0.0065	build	0.0068
impact	0.0064	effect	0.0062	help	0.0064
fund	0.0063	storm	0.0059	support	0.0063

Table 2: Topics and top twenty words and probabilities per topic

From the most frequent words in each topic, we can already sense that certain countries are more dominant within specific topics. The US and its current president features most prominently in the ‘politics’ topic, while China is salient within the ‘development’ topic. Later, through semantic analysis, we will explore whether this may be largely due to the self-referencing (do the US and China refer to itself frequently?) or references from other countries (do other countries refer to these countries, too?).

Through topic modeling, we found that countries paid unequal attention to the three topics. All countries allotted some attention to the ‘political’ part of climate change, though developed and developing countries diverged in the other two topics: developed countries generally focused on the ‘environment’ aspect of climate change, while developing countries tended to focus on economic ‘development’. We assessed this by using the same model and calculating the estimated proportion of each document corresponding to each topic. Then, the average proportions allotted to each topic was calculated for each country (Table 3). We found that all eight countries devoted a significant amount of article space to the politics of climate change (Figure 4). This proportion ranged from a low of about 25 - 30% in Kenya, Australia, China, and the Middle East to a high of almost half in India (48%)

and the UK (44%). This suggests that countries of all types recognize that climate change is part of an international, political dialogue that no country can unilaterally ignore.

However, countries showed much more variance in the proportion of article-space devoted to the environmental and scientific aspects of climate change. Nigeria and China showed the least interest in the environmental sphere of climate change (14% and 17%, respectively), while Australia and the US were at the other end of the spectrum with each devoting about 60% to the topic. Likewise, countries focused in varying degrees to the economic development aspect of climate change. Australia, the UK, and the US showed little interest in the economic development (around or less than 10% each), while Kenya and Nigeria each devoted over half their article-space to the topic. Particularly interesting is that at 54%, China focused the most attention to this topic. We can see a clear divide between more developed nations (Australia, UK, and US) focusing on the environment, and developing nations (China, India, Kenya, Nigeria) focusing on the economic development challenges related to climate change. This also suggests that China, partnering with developing countries, is poised to play a leadership role in the economic development aspect of climate change. The Middle East is most balanced in topic proportions, though it is somewhat skewed towards development rather than environment.

Country	Politics	Environment	Development
Australia	0.2590	0.6199	0.1194
China	0.2825	0.1744	0.5416
India	0.4805	0.2761	0.2415
Kenya	0.2545	0.2249	0.5190
Middle East	0.2968	0.2899	0.4117
Nigeria	0.3402	0.1371	0.5208
UK	0.4355	0.5043	0.0586
US	0.3686	0.5772	0.0529

Table 3: Average proportions for each topic by country

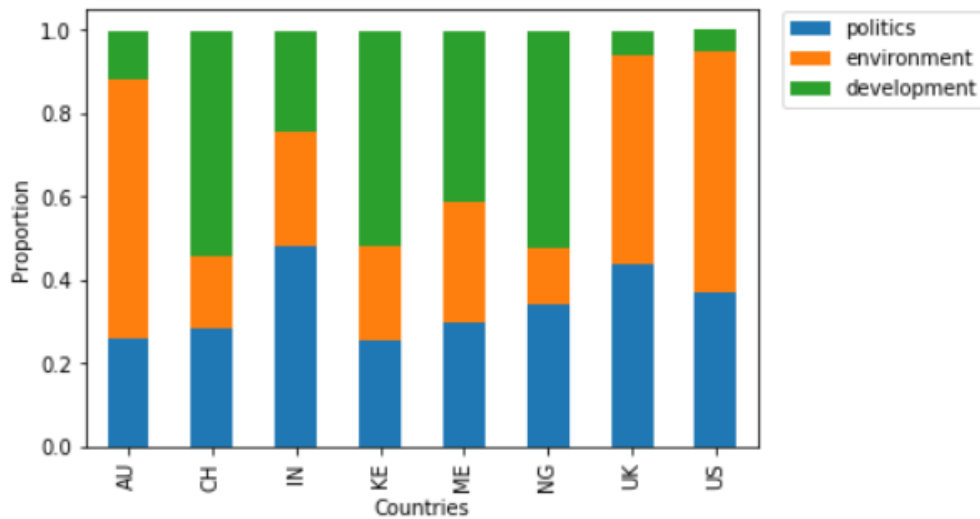


Figure 4: Average proportions for each topic by country

We checked the stability of country-wise topic proportions by aggregating in a different way. Instead of averaging the proportions across all documents in each country (above), each document is assigned its top topic exclusively and the document proportions were then calculated for each country. For example, if the proportions for the politics and development topics were 33% each and the environment topic proportion was 34%, then the document was classified as an environment document (in reality, such even distributions of topics were rare). This aggregation produced almost the same distribution of topics within every country (Table 4). This suggests that each country’s focus in each of the topics is very stable (Figures 4 and 5 are nearly identical).

Country	Politics	Environment	Development
Australia	0.2515	0.6535	0.0950
China	0.2664	0.1768	0.5568
India	0.4915	0.2812	0.2273
Kenya	0.2450	0.2138	0.5412
Middle East	0.3036	0.2883	0.4081
Nigeria	0.3232	0.1443	0.5325
UK	0.4358	0.5229	0.0413
US	0.3442	0.6085	0.0472

Table 4: Proportion of documents for each topic by country

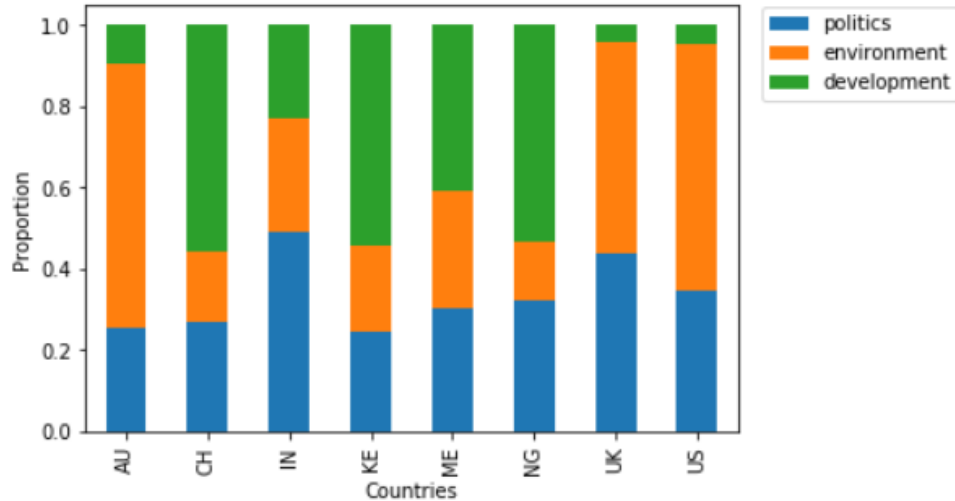
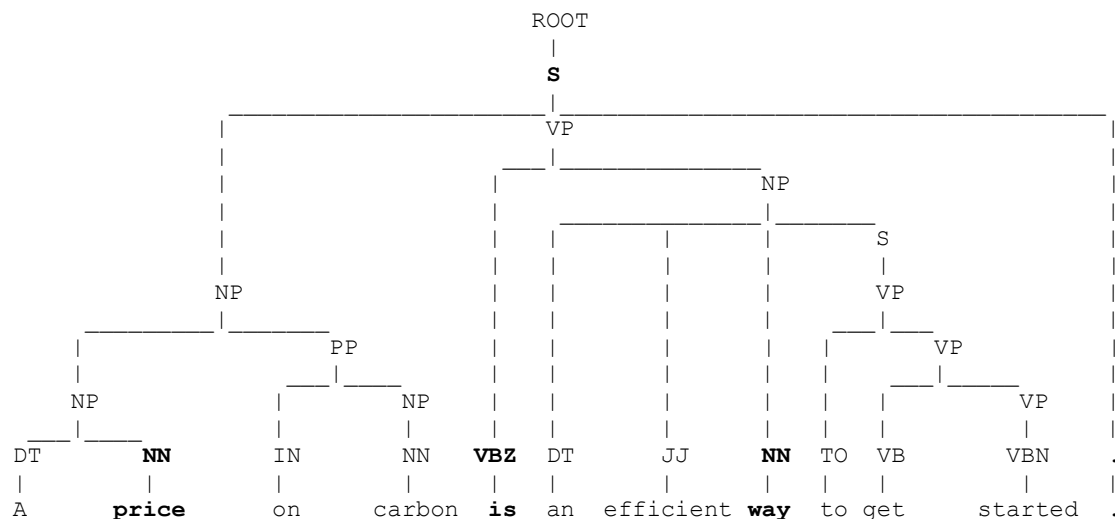


Figure 5: Proportion of documents for each topic by country

V. Semantics

We have previously used approaches based on word embeddings and topic modeling to explore the aggregated meaning behind the news articles. Through semantic analysis, we will look at the deeper meanings and connections present in the corpus. This entails looking at the sentences in the text and analyzing which words fill which roles in the sentence, for instance nouns, verbs, adjectives that in turn act as direct objects, subjects, etc. Using the Natural Language Toolkit for Python and the Stanford Parser, words were tagged with the “part-of-speech” they fulfill in the sentence. A dependency tree that represents the relationship between the grammatical components of sentences

was built for each of the sentences in the articles. It should be noted that, as with all computational modeling inferences, Stanford parser was not 100% reliable, especially for complex sentences, but in general yielded reasonable results. Due to constraints in computational power⁷ that prevented us from applying this to our entire corpus, we randomly sampled 3000 articles from the original sample of 6000 articles.



Once the sentences were parsed, we looked at actors, actions, and receivers of the actions, or Subject-Action-Object (SAO) semantic triples, in the news to understand what was happening in each country (Franzosi 1994). To aggregate the SAO triples, we selected only simple declarative sentences, that is, sentences that don’t have subordinate clauses replacing either the subject or object. To further aid aggregation, we selected only the last word from a multi-word noun (so that “Paris Agreement” is only referred to as “agreement”). For the verbs, we removed auxiliary verbs (such as ‘did’, ‘have’, ‘been’) from multi-word verbs, and lemmatized the verb to its infinitive form to aid with aggregation. As an example, for the main sentence in Figure 6, the parsed SAO triple is “price be way”.

A. United States of America

Subject	Action	Object	Count	Top Subjects	Top Actions
change	be	hoax	27	change	be
change	be	issue	18	i	have
change	be	threat	14	you	say
change	be	problem	12	trump	make
climate	change	hoax	11	people	take
activities	be	cause	10	administration	cause
share	be	parts	10	scientists	tell
change	be	driver	10	climate	s
change	increase	risk	10	warming	change
deniers	be	equivalent	9	humans	affect
evidence	abound	top	9	report	do
house	release	report	8	states	come

⁷ Machines utilized had a hard limit on RAM usage, as part of the University of Chicago Research Computing Cluster.

change	cause	increase	8	t	increase
businesses	take	steps	8	temperatures	get
dioxide	be	contributor	8	percent	become
times	publish	report	8	changes	contribute
administration	release	report	8	students	find
facts	be	facts	8	countries	see
change	be	oceans	7	pruitt	include
references	remove	websites	7	them	call

Table 5: Most common Subject-Action-Object combinations, Subjects, and Verbs for the US

The picture presented by the most common SAO combinations in the United States is that of a country divided. Climate change-denying and affirming views are reflected in the reporting in the news and in our analysis, perhaps most tellingly demonstrated by “[climate] change [is] hoax” being the most repeated simple declarative sentence. The opposing view is represented by the triples “[climate] change cause increase”, “facts be facts”, “[climate] change is driver”, among others (Table 5). These contrasting positions dominate the top triples, reflecting the US’s overall environmental topic focus. The country’s secondary focus on politically-related climate change reporting, as seen previously in the topic modeling section, is cemented by “Trump” being one of the most common sentence subjects. This is not reflected in the most common SAO triples because of the diversity of actions and objects that “Trump” the subject takes on.

B. United Kingdom

Subject	Action	Object	Count	Top Subjects	Top Actions
climate	change	hoax	6	trump	be
uk	remain	agreement	6	change	have
i	make	position	5	i	make
decision	draw	condemnation	5	us	take
change	lead	rises	4	you	say
world	implement	agreement	4	people	s
may	express stress	disappointment	4	countries	tell
you	become	hero	4	government	come
accord	commit	countries	4	investors	affect
scientists	warn	failure	4	companies	see
films	play	part	4	scotland	give
world	be	mess	4	world	put
change	be	issue	4	decision	do
shift	be	technology	4	agreement	include
change	be	problem	4	climate	become
tillerson	tell	senators	3	leaders	lead
corbyn	label accuse	move	3	accord	meet
us	withdraw	accord	3	group	call
us	need	seat	3	may	cause
us	give	notice	3	plan	change

Table 6: Most common Subject-Action-Object combinations, Subjects, and Verbs for the UK

Climate change news in the UK focuses heavily on aspects of international politics, including the UK's own stance regarding events in this arena, and also remarks on the environmental effects of climate change (Table 6). Top among the SAO triples are combinations like “decision draw condemnation”, “world implement agreement”, and “accord commit countries”. These clearly refers to the Paris Agreement, which requires international commitment and elicits censure for those countries that do not cooperate. The UK also reiterates its own stance in the international politics of climate change; Theresa “may express[es] disappointment” [about the US decision to withdraw from the Paris Agreement], and the “uk [will] remain [in the] agreement”. The UK's focus on politics is also reflected in the most frequent subjects, which include words like ‘countries’, ‘government’, and ‘agreement’. In addition, the most frequent verbs are more descriptive and not very consequential, reflecting a political scene of concepts rather than actions. There is also some mention of the effects of climate change on the environment through SAO triples like “climate change hoax”, “change lead [to temperature or sea level] rises”, and “scientists warn failure”.

C. Australia

Subject	Action	Object	Count	Top Subjects	Top Actions
change	be	problem	8	change	be
change	be	issue	8	i	have
businesses	have	freedom	6	you	s
change	happen	change	6	government	make
communities	do	lifting	6	australia	take
us	change	model	6	people	say
people	kill	goats	6	t	tell
move	stifle	science	6	us	do
remediation	have	value	6	trump	come
reform	represent	charities	6	abbott	put
reform	represent	challenge	6	companies	become
activity	lead	prosperity	6	council	go
millions	spend	business	6	resolutions	increase
donations	rise	years	6	turnbull	cause
services	rely	data	6	scientists	change
tanks	drive	agenda	6	report	see
council	set	mission	6	climate	call
resolutions	win	support	6	warming	include
businesses	use	clout	6	temperatures	lead
warming	replace	change	5	policy	provide

Table 7: Most common Subject-Action-Object combinations, Subjects, and Verbs for Australia

The most common SVO combinations in Australia's news, in contrast to the US, do not present a picture of doubt on whether climate change is real (Table 7). Climate change is presented as a “problem”, an “issue”, and something that “happen[s]”. Australian news discusses climate change as something that is occurring and seems to have the view that is friendlier towards scientific approaches; it often mentions “move[s that] stifle science” and suggests the need to do something about it, as exemplified by “remediation [has] value” and “reform[s] represent challenge”. The climate change

discourse in Australia also seems to have a business-friendly slant, as “businesses have freedom” and “use clout”.

D. China

Subject	Action	Object	Count	Top Subjects	Top Actions
china	play	role	19	china	be
china	be	market	6	countries	have
china	be	country	6	trump	say
trump	announce	withdrawal	5	i	make
trump	sign	order	5	us	take
economy	grow	percent	5	xi	tell
daily	ask	experts	5	people	become
china	be	partner	5	author	hold
trump	take	office	5	you	come
china	become	partner	4	government	play
china	become	economy	4	j Jinping	give
china	make	progress	4	country	see
party	hold	Congress	4	initiative	provide
china	be	economy	4	economy	bring
landslide	kill	people	4	leaders	show
china	revise	law	4	world	set
eu	be	partner	3	change	call
speakers	discuss	change	3	cooperation	face
daily	condense	article	3	meeting	sign
china	contribute	percent	3	states	work

Table 8: Most common Subject-Action-Object combinations, Subjects, and Verbs for China

China paints a very distinctive picture of its role in the climate change scene – it is very keen on playing a central role in addressing climate change through economic development (Table 8). By far the most common SAO set is “china play role”, hinting at its desire to fill the climate change leadership role that was left open when “trump announce[d] withdrawal”, another top triple. In particular, China clearly wants to play a key role in economic development, as evidenced by such SAO combinations like “china [is] market”, “economy grow percent”, and “china become economy”. It is also very keen on forging international partnerships as indicated by “china be partner” and “eu become partner”.⁸ Common subjects like ‘initiative’, ‘economy’, and ‘cooperation’, which do not occur often in other countries, and verbs, including ‘play’, ‘give’, ‘provide’, and ‘bring’, also show its development-oriented economic focus.

⁸ Some of the most common SAO word sets were about authors, such as (‘author’, ‘be’, ‘researcher’). This shows the authority of the writers of the articles, and probably reflects China’s respect for, rather than skepticism of, expert opinion. However, these were excluded in the analysis because it does not tell much about China’s attitude toward climate change itself.

E. India

Subject	Action	Object	Count	Top Subjects	Top Actions
writer	be	journalist	9	india	be
change	be	issue	7	countries	have
ministry	change	name	6	change	take
research	be	tune	6	trump	make
zsi	collaborate	survey	6	us	say
zsi	set	system	6	i	lead
india	double	production	6	government	come
cover	study	impacts	6	people	give
temperature	increase	celsius	5	agreement	tell
islands	identify	setting	5	you	become
byravan	be	scientist	5	world	increase
work	start	islands	4	study	affect
change	be	challenge	4	china	provide
us	pull	agreement	4	nations	see
actions	reduce adapt	emissions	4	country	do
us	be	emitter	4	farmers	set
countries	ratify	amendment	4	modi	hold
countries	take	targets	4	report	include
sharma	be	berlin	4	states	cause
change	combat	adaptation	4	decision	go

Table 9: Most common Subject-Action-Object combinations, Subjects, and Verbs for India

India voices diverse views on climate change, ranging all three of the topic spaces (Table 9). India recognizes the environmental impact of climate change and the role of scientific research through top SAO triples that include “change [is] issue”, “research [is] [in] tune”, and “temperature increase celsius”. India also covers the international political aspects of climate change, focusing particularly on agreements on [carbon] emissions through the following SAOs: “us pull agreement”, and “us [is] emitter”, and “countries take targets”. Economic development is also featured through the SAO triple “india double production”. Interestingly, ‘china’ is one of the most frequent subjects in Indian news, indicating China’s prominent geo-political position in the region. In aggregate, India’s portrayal of climate change seems to be one of the most variegated among the countries analyzed.

F. Nigeria

Subject	Action	Object	Count	Top Subjects	Top Actions
i	tell	you	8	i	be
nigeria	participate	meetings	5	you	have
service	record	collection	5	nigeria	make
buhari	appeal	community	4	government	say
hands	be	deck	4	people	s
i	assure	you	4	change	take
nigeria	ratify	agreement	4	trump	give
buhari	sign	agreement	4	buhari	come
inflation	fell	months	4	countries	do

officers	wait	pensions	3	country	tell
me	tell	you	3	president	include
country	be	illusion	3	minister	urge
eurobonds	raise	billions	3	project	call
nigeria	implement	contributions	3	us	provide
you	get	response	3	plan	put
bonds	be	instruments	3	state	go
delegation	canvas	support	3	world	become
fight	take	turn	3	bank	get
world	face	today	3	states	hold
aircraft	land	airport	3	initiative	use

Table 10: Most common Subject-Action-Object combinations, Subjects, and Verbs for Nigeria

Nigeria focuses extensively on the international political aspect of climate change, in addition to the economic development aspect (Table 10). The most common SAO triples include “nigeria participate [in] meetings”, “nigeria ratify agreement”, “[president] buhari appeal[s] [to] community”, and “buhari sign[s] agreement”. However, in contrast to other countries, Nigeria seems to be more of a receiver instead of an initiator of action in the international politics scene. We can sense this in SAOs such as “delegation canvas support”, and in common verbs such as “urge” and “call”, which suggests the need for persuasion rather than decisive action. Another focus in Nigeria is economic development, which is evidenced by SAO combinations like “inflation fell months” and “bonds be instruments”. This is also reflected in the most common subjects, which include “minister”, “project”, “plan”, and “bank”.

G. Kenya

Subject	Action	Object	Count	Top Subjects	Top Actions
kenya	ratify	agreement	4	i	be
kenya	be	countries	4	kenya	have
i	tell	you	4	you	s
kenya	have	potential	3	farmers	take
landslide	kill	people	3	people	say
drought	experience	region	3	trump	make
river	be	resource	3	government	tell
change	be	threat	3	countries	face
states	roll	actions	2	change	come
government	rule	link	2	africa	provide
county	partner	industries	2	obama	use
county	head	crisis	2	country	become
i	manage	bags	2	them	call
county	realise fear	bags	2	project	go
massacre	fall	victim	2	drought	do
people	be	risk	2	us	experience
change	generate	dynamics	2	governments	affect
immelt	write	twitter	2	sector	lead
awards	present	individuals	2	t	include
i	know	power	2	company	see

Table 7: Most common Subject-Action-Object combinations, Subjects, and Verbs for Kenya

Kenya’s most common triples paint a dire picture of the environmental consequences of climate change (Table 11). Extremely negative SAO triples like “landslides kill people”, “drought experience[d] [by] region”, “people be [at] risk”, and “[climate] change be threat” possibly reflect the concerns of a developing country trying to manage the environmental impacts of climate change with little resources. Kenya’s development topical weight is represented by discussions of Kenya as “having potential”, and “river be resource”, in addition to “farmers” counting as one of its most common sentence subjects (even ahead of “Trump”). Notable among the most commonly used verbs in Kenya’s news items are “call”, as in the case of calling or urging for cooperation or assistance.

H. Middle Eastern Countries

Subject	Action	Object	Count	Top Subjects	Top Actions
project	be	part	5	i	be
people	mistake	change	5	you	have
change	impact	economy	5	trump	s
kaabi	attend	ceremony	4	people	take
ministry	say	statement	4	countries	tell
mansoori	attend	ceremony	4	change	say
change	be	issue	4	us	make
policymakers	work	summit	4	t	become
attack	show	contempt	3	uae	come
part	s	week	3	government	use
footage	post	media	3	ministry	see
people	live	lives	3	world	include
mills	be	ceo	3	china	play
paper	say	editorial	3	companies	do
agreement	come	force	3	country	hold
cities	have	role	3	egypt	lead
events	hold	conjunction	3	leaders	face
europe	be	counterpart	3	states	announce
i	don	t	3	one	give
spokesman	call	release	3	technology	need

Table 8: Most common Subject-Action-Object combinations, Subjects, and Verbs for Middle Eastern nations

The combination of Middle Eastern Countries (Egypt, UAE and Saudi Arabia) produced a set of SAO combinations that do not seem to have a particularly strong direction or focus (Table 12). One of the most interesting sentences is that “[climate] impact[s the] economy”, reflecting a dimension of climate change that is important to a region heavily reliant on fossil fuels. Otherwise, the Middle East tended to focus on the more official or ceremonial aspects of government through SAO triples like “kaabi attend ceremony”, “ministry say statement”, and “policymakers work summit”.

Through semantic analysis we saw that all countries acknowledge that climate change is part of an international, political exchange where countries themselves are actors, constantly responding to and negotiating with other actors. Further, some countries (such as China and the UK) seem to be actively trying to situate itself in this scene through the news medium, while others (such as Nigeria) use the medium as a forum to urge cooperative action. In addition, we can see that in most countries, there is

a recognition of the negative effects that climate change has on the environment, with many countries also acknowledging the links to science.

It is worth noting that some of the most common subjects in all countries across the board are “Trump” and “US”. This is probably at least in part due to the controversial US withdrawal from the Paris Agreement in June 2017. However, it also reflects the prominent position the US occupies in world politics (another country’s withdrawal may not have received so many mentions). This also shows that “Trump” and the “US” being the most probable words in the politics topic is probably due to all countries referring to them in high frequency. In addition, the self-referencing of countries to itself was a very common phenomenon across countries. China by far referenced itself the most, with all other subjects lagging far behind. Aside from India, no other country mentioned China frequently, suggesting that China’s prominence in the economic development topic is probably due to self-referencing. The word “people” also appears as one of the most frequent subjects in all countries. “People” can be used in many contexts, but its frequency may also reflect the wide-ranging effect climate change has on people everywhere.

VI. Discussion

Do countries with similar economies or institutions produce similar news items regarding climate change? In this section, we will explore some of the general patterns between countries’ characteristics and their attitude toward climate change.

The most powerful trend we found was the tendency in high-income countries to focus much attention on the environmental aspects of climate change, while lower-income countries tend to focus on economic development. Table 13 shows the per capita GNI (World Bank 2018) of the countries we examined, juxtaposed with the proportion of news covering environmental and economic development topics. We can see a clear positive relationship between GNI and coverage of environmental topics, and a negative relationship with development. Having established, developed economies, wealthier countries may have more resources and bandwidth to focus on the environmental effects of climate change. China is in the middle ground between developed and developing countries in terms of GNI, but clearly identifies with the developing countries in terms of its emphasis on economic growth within the climate change context. Oil-rich Middle Eastern countries may be concerned about future growth prospects with decreasing demand or sources of fossil fuel. The focus on economic development in countries like Kenya and Nigeria may reflect their dual concern of acquiring resources like water and food and achieving sustainable growth in a world impacted by adverse climate. The resources a country has is clearly related how climate change is covered in the country.

Country	GNI per capita, 2016, current USD	Environment	Development
United States	\$58,700	0.5772	0.0529
United Kingdom	\$41,640	0.5043	0.0586
Australia	\$45,210	0.6199	0.1194
Egypt, UAE, Saudi Arabia (weighted avg.)	\$25,687	0.2899	0.4117
China	\$15,470	0.1744	0.5416
India	\$6,490	0.2761	0.2415
Nigeria	\$5,740	0.1371	0.5208
Kenya	\$3,120	0.2249	0.5190

Table 13: 2016 GNI per capita (PPP) in current USD, topic proportions for environment and development

It should also be noted that one of the muddling factors of the analysis is the fact that not all countries present a similar degree of press freedom. Press freedom can affect the degree to which news items produced in a country are an accurate reflection of the nation's citizens views and attitudes, instead showing the state's – or the power holder's. Table 14 shows Reporters Without Borders' Freedom of the Press Index for 2017 (Reporters Without Borders 2018). Looking at the Index in conjunction with our results, there seems to be a general positive correlation between free speech and the degree to which countries report news in the environment topic; that is, countries with a higher degree of press freedom (a lower index) devote a higher proportion of its news coverage to environmental news. A potential explanation for this is that reporters in these countries have more flexibility to assess their nation's role in contributing to climate change and are free to publish critical articles. They may also have more autonomy in choosing which topics to report, reflecting public sentiment rather than state support. This is a very general observation that merits further analysis in a separate study.

Country	Freedom of the press index, 2017
United States	23.88
United Kingdom	22.26
Australia	16.02
Egypt, UAE, Saudi Arabia (avg.)	53.30
China	77.76
India	42.94
Nigeria	39.69
Kenya	31.20

Table 14: 2017 Freedom of the press index

Surprisingly, there does not seem to be a clear trend between per capita carbon emissions and the way in which climate change is covered. Table 15 shows the per capita carbon emissions by country (Emissions Database for Global Atmospheric Research 2018). Some countries with higher emissions rates (such as Australia) focus heavily on environmental topics, as do some countries with a lower carbon footprint (such as the UK). Carbon emissions is the result of a complex set of factors, including a country's geography, industry, etc., requiring more nuanced analysis. Carbon emissions is more

difficult to correlate with climate change news coverage, and more countries may need to be analyzed at more depth for meaningful findings.

Country	Carbon emissions (ton) per capita, 2016
United States	15.56
United Kingdom	5.59
Australia	17.22
Egypt, UAE, Saudi Arabia (weighted avg.)	6.96
China	7.45
India	1.92
Nigeria	0.44
Kenya	0.34

Table 159: 2016 carbon emission per capita in tons

The unobserved, “cultural” characteristics of a country probably has the greatest impact on how climate change is covered in the news. Across the different methods we used, the United States seems to be unique. The debate about climate change denialism seems localized to the United States, where it consumes a big part of the climate change news conversation, as seen in the SAO triples. Additionally, mapping “America” on our “veracity” dimension places it alone on the end of the spectrum that is filled with doubt about the scientific consensus. This is likely the reason its news articles also possess a slightly negative mean similarity to most other countries. The United States is alone having this debate, while the world moves on. Contrastingly, it seems China is wanting to fill the power void left in this space. Chinese news items clearly paint China as a leader, as evidenced by the prevalence of leadership-related SAO triples and its placement on the leadership dimension of word embeddings. Moreover, China’s topical focus on development and SAO triples on the “economy” show how it intends to manifest this role.

VII. Conclusion

Using natural language processing methods on a large sample of newspaper articles revealed that countries approach climate change in varied ways. Word embeddings showed general similarities among countries, and highlighted how the US stands out in its difference. Topic modeling revealed the broad themes within the climate change space, and further showed that all countries devote some attention to the politics of climate change. Semantic analysis shed deeper insight into who is doing what (and to whom). Interestingly, countries’ attitudes were somewhat linked to their income levels, with wealthier countries focusing on the environmental aspects and lower income countries focusing on the economic development aspect of climate change. Notably, the American environmental focus is largely dominated by the debate on the veracity of climate change. On the other hand, China seems poised to lead the economic development aspect of climate change, and is ready to partner with other countries. We have delved into some of the different perspectives that countries have on climate change. Deeper understanding of where and how countries agree and diverge will help the international community to arrive at consensus that galvanizes action.

Bibliography

- Blei, David M. 2012. "Probabilistic Topic Models." *Communications of the ACM* 55 (4).
- Caliskan, et al. 2017. "Semantics Derived Automatically from Language Corpora Contain Human-like Biases." *Science* (356): 183-187.
- Emissions Database for Global Atmospheric Research. 2018. *Fossil CO₂ & GHG emissions of all world countries, 2017*. <http://edgar.jrc.ec.europa.eu/overview.php?v=CO2andGHG1970-2016&sort=des8>.
- Franzosi, Roberto. 1994. "From Words to Numbers: A Set Theory Framework for the Collection, Organization, and Analysis of Narrative Data." *Sociological Methodology* 24: 105-136.
- Le, Quoc. 2014. *Distributed Representations of Sentences and Documents*. <https://arxiv.org/abs/1405.4053>.
- Reporters Without Borders. 2018. *World Press Freedom Index*. <https://rsf.org/en/ranking/2017>.
- World Bank. 2018. *GNI per capita, PPP (current international USD)*. data.worldbank.org/indicator.