WWW.JBDTP.ORG

ISSN: 2692-797

JBDTP Professional Vol. 1, No. 1, 2022 DOI: 10.54116/jbdtp.v1i1.18

GLOBAL PERCEPTION OF THE BELT AND ROAD INITIATIVE: A NATURAL LANGUAGE PROCESSING APPROACH

Amit Mokashi
New Jersey City University,
School of Business
amokashi@njcu.edu

J.D. Jayaraman
New Jersey City University,
School of Business
jjayaraman@njcu.edu

Priyanka Mahakul
New Jersey City University,
School of Business
pmahakul@njcu.edu

Rutu Patel
New Jersey City University,
School of Business
rpatel19@njcu.edu

Anthony Picciano
New Jersey City University,
School of Business
apicciano@njcu.edu

ABSTRACT

In fewer than seven years since the launch of the Belt and Road Initiative, 138 countries have signed onto the program, with, by some counts, 118 projects being planned. The Belt and Road Initiative is a Chinese multi-trillion–dollar global infrastructure initiative that has geopolitical implications for both participating and nonparticipating countries. Some of the unique selling points of this initiative also make it controversial among its stakeholders. These variations in sentiments can be perceived in the media reporting in which there is freedom of expression. This paper uses sentiment analysis to gauge the variation in the stakeholder perception over time across three groups: China, participating countries, and nonparticipating countries. Our analysis of 7,856 news articles has provided quantitative evidence of declining positive sentiment over time.

Keywords *Belt and Road Initiative, One Belt One Road, Natural Language Processing, Sentiment Analysis, Opinion Mining*

1. Introduction

1.1 Historical Context

President Xi Jin Ping of China in 2013 announced what was then known as the One Belt One Road strategy (Chatzky and McBride 2019). The belt in this strategy referred to the terrestrial corridors, whereas the road referred

to the maritime lanes. This plan was a modern take on the ancient Silk Route/Road. The original Silk Road was not singular, nor did it only facilitate the silk trade. It was a network of roads that carried different goods from different countries. From West to East, these goods included horses, saddles and riding tack, grapevine and grapes, dogs and other animals both exotic and domestic, animal furs and skins, honey, fruits, glassware, woolen blankets, rugs, carpets, textiles (such as curtains), gold and silver, camels, slaves, and weapons and armor. From East to West, the goods included the following: silk, tea, dyes, precious stones, china (plates, bowls, cups, vases), porcelain, spices (such as cinnamon and ginger), bronze and gold artifacts, medicine, perfumes, ivory, rice, paper, and gunpowder (Mark 2019). The road carried more than goods. It also brought migrants, religion, science, and art (UNESCO 2019). Its amalgamation, therefore, was very organic. In fact, the network was not called Silk Road until 1877 (Whitfield 2007). As with the original Silk Road, the modern version is not a single road but a network of roads. The name, therefore, was subsequently changed from One Belt One Road to the Belt and Road Initiative (BRI) to reflect this broader scope (Bērziņa-Čerenkova 2016).

1.2 China's Plans for Its New Silk Road

The idea of rejuvenating the ancient Silk Road is not new and has been discussed and advocated enthusiastically in the past (Griffiths 2017). However, the magnitude of China's approach seems to have created apprehension in policymakers adversely impacted by the project. President Xi plans to develop a network of roads, railways, pipelines, and ports to facilitate trade with the world. China's investment in this global initiative is expected to cross well over a trillion dollars. The BRI has six main economic corridors: (1) the New Eurasian Land Bridge, (2) the China-Central Asia-West Asia Corridor, (3) the China-Pakistan Corridor, (4) the Bangladesh-China-Myanmar Corridor, (5) the China-Mongolia-Russia Corridor, and (6) the China-Indochina Peninsula Corridor (Indermit Gill and Mathilde 2019). China claims that this project is economic in nature and would be a mutual win-win for all the parties involved. Many countries, however, have been skeptical about both the intent as well as the consequences of this mega project on the host nations (Chatzky and McBride 2019).

1.3 Concerns about the BRI

There are two main categories of apprehensions with regard to this initiative: economic and strategic. Economically, the problems revolve around debt implications for the countries taking on the BRI-linked projects. At the same time, strategically, the concerns are about the implication of the indebtedness of the borrowers on their ability to make independent policy decisions (Hurley, Morris, and Portelance 2019). An excellent example of this is Sri Lanka's Hambantota port project (Rithmire and Li 2019). Sri Lanka took a loan of more than a billion dollars from the Chinese Ex-Im bank to pay for the Chinese-built port in southern Sri Lanka. The port, however, did not generate the projected revenue, and the Sri Lankan government ended up handing over the port in lieu to China. This concern of the borrowing countries has led to public opposition to the BRI-funded projects in Sri Lanka, Maldives, Malaysia, Kenya, and Pakistan (Balding 2018). Other countries, such as the United States of America and India, have reservations due to the strategic implications of the growing Chinese influence on the BRI-participating countries. Indeed, infrastructure that is used to support trade can be equally efficiently used to support the military. The Chinese have been accused of being opaque in their dealings and resorting to bribery to get the decisions in their favor. China has made conscious efforts to overcome this image by both being open in its dealings (Bloomberg 2019) and disseminating information (Adrien 2019). China has renegotiated some of its projects while it has also written off some of the loans that it has given. These steps have had some success in resurrecting stalled projects and public opinion.

Anecdotal evidence shows that the perception about the BRI has been swinging both in the positive as well as the negative direction. An understanding of the public perception of the BRI has policy implications in participating and nonparticipating countries. Thus, it is important to understand the public sentiment of the BRI in a systematic and unbiased manner. In this paper, we explore public sentiment of the BRI in a systematic and unbiased manner by using a natural language processing (NLP) based approach. NLP techniques allow us to analyze large amounts of textual data in an automated fashion, thereby enabling us to extract useful information from vast quantities of documents in a cost-effective and unbiased manner.

Without the use of NLP, a large-scale analysis such as the one presented in this article will not be practically feasible because the alternative would be for a human to manually analyze the documents. The magnitude and direction of public opinion in both temporal and spatial domains are examined by using sentiment analysis, an NLP technique. This is among the first applications of an NLP-based approach to investigating public opinion about the BRI and is one of the few applications of NLP in the field of transportation in general. We assessed public opinion on the BRI from news articles in the international media. The use of sentiment analysis overcomes the challenge of manually identifying, aggregating, and summing diverse opinions and trends from a large dataset of more than 7,000 news articles published over an approximately five-year period. The automated sentiment analysis process

also removes any human bias in determining opinion. Thus, our paper contributes to the meager extant literature on NLP applications to transportation by introducing a novel method of determining public opinion on one of the major global initiatives with significant geopolitical importance.

2. Literature Review

Research on the application of machine learning and artificial intelligence (AI) techniques in transportation is still in its infancy, although substantial progress has been made in recent years. There is limited extant literature that we are aware of applying NLP techniques to assess public opinion of the BRI. The earliest research articles that applied sentiment analysis in the BRI context seem to have been published around 2019. Niu and Wu (2019) used the "Belt and Road" related corpus to evaluate different sentiment analysis methods. They found that the accuracy of the dictionary-based method depended on the comprehensiveness of the selected emotional dictionary. Arifon *et al.* (2019) compared the Chinese and European discourses with regard to the "Belt and Road Initiative." They found the European media suspicious of the Chinese media, which they considered a voice of the Chinese ruling party. Li *et al.* (2019) ranked the consumer's risk perception on nine BRI countries from high to low: Czech Republic, Thailand, Malaysia, Turkey, Hungary, Poland, Russia, Singapore, and Romania. Ali *et al.* (2020) conducted a thematic content analysis of opinion in the Pakistani Twittersphere. They found that the technicians of opinion were effectively adopting the multi-thematic discourse and portraying the China-Pakistan Economic Corridor (part of the BRI) as a landmark project.

A systematic review and comparative assessment of the Chinese and English-language literature on the environmental impacts of China's BRI (Teo *et al.* 2020) found that much of the Chinese literature may be targeted for domestic consumption and thus may not contribute to the international discourse on BRI. A paper by Chandra *et al.* (2020) that aimed to extract the sentiments of the BRI from Twitter found positive sentiments dominant among the extracted tweets. However, closer inspection revealed that some of those classed positive tweets were, in fact, sarcastic in nature. Malik (2020) used rhetorical theory to analyze the Chinese official report in 2019, the American versus European media response to the BRI project, and the US direct response to the BRI in the Indo-Pacific Strategy in 2019. Napitupulu *et al.* (2020) analyzed public sentiment through social media and Twitter on foreign workers in Indonesia during the coronavirus disease 2019 (COVID-19) pandemic. The data were collected by using social network analysis by identifying related topics in Drone Emprit Academic (software for social media monitoring and analytics).

The study results indicated negative sentiments toward the existence of foreign workers, especially those from China. Li *et al.* (2021) developed a novel socio-environmental sensing approach by synthesizing remote sensing imagery and geotagged Twitter data to map the socio-environmental impact of Large-scale infrastructure projects. Their focus was on two BRI flagship projects, namely, the Mombasa-Nairobi Standard Gauge Railway in Kenya and the China-Pakistan Economic Corridor in Pakistan. In that context, they found that public sentiment toward the projects was largely positive and improved over time. Zhou *et al.* (2021) conducted a public opinion analysis of a very specific component of the BRI, that is, the Linyi trade service-oriented country logistics hub in Shandong, China. Their scope was limited to contents in CCTV1, People's Daily, People.com.cn, Guangming.com, Chinanenews.com, Ministry of Commerce, China Media Group, and CnR.cn, where they found the opinions to be largely positive.

Few research articles have applied NLP techniques to assess passenger satisfaction, service quality, and community engagement. The most common NLP technique used in these studies is sentiment analysis, which examines language in social media and blogs to detect positive or negative emotion words and words that convey, for example, satisfaction, engagement. Liu, Li, and Li (2019) used sentiment analysis to investigate the public transportation comments found on the Dazhong-Dianping Shanghai Station website to extract opinions and to determine transportation service satisfaction. Evans-Cowley and Griffin (2012) used sentiment analysis to investigate microparticipation in the Austin Strategic Mobility Plan. They analyzed 49,000 posts on Twitter and Facebook to determine public engagement with the strategic planning process. Das *et al.* (2018) investigated Twitter data with bike commuting hashtags to understand factors that influence people to bike commute. They used exploratory text mining and sentiment analysis to analyze how people's opinion on bike commuting has changed over the years.

Apart from sentiment analysis, some studies used text mining to analyze unstructured textual data. Gu, Qian, and Chen (2016) mined Twitter data to detect traffic incidents in real time. They used keywords and their embeddings to train a classifier to determine whether there was a traffic incident. Mehrotra and Roberts (2018) analyzed the vehicle owner's questionnaire crash data collected by the National Highway Transportation and Safety Administration by using latent semantic analysis to identify the emergent themes that capture the key issues that vehicle owners encountered. Thus, research that applied NLP techniques to unstructured data to analyze transportation-related problems is meager. We contribute to this scant extant literature by applying NLP techniques to analyze public opinion about the BRI with important geopolitical ramifications.

3. Research Questions

The research questions we investigated in our study are as follows:

1. What is the overall public sentiment about the BRI gathered from the international news media?
2. How does this sentiment change over time?
3. How does this sentiment break down by categories (China, BRI participants, and BRI nonparticipants)?
4. What are some key drivers of the changes in BRI-related sentiment over time across the clusters?

4. Methodology

4.1 Data

The data consisted of news articles in English about the BRI. Using only the English language articles made it possible to use a common lexicon. These articles were collected from the Factiva database. Factiva is an international news database curated by Dow Jones, New York, USA, a leading provider of news and financial information. It offers a powerful research platform, with comprehensive coverage of international news from more than 32,000 sources. We searched the Factiva database for the keywords "Belt and Road Initiative" over the time period from the BRI initiative announcement in September 2015 to May 2020, which resulted in 7,856 news articles. Many of these were duplicates, that is, original opinions republished by other sources. To get a more realistic idea of the source of the sentiment, the republished articles were ignored and the original agency and its location were credited as being the source. This brought the total number of unique articles down to 1,419 and their sources to 57 countries. The distribution of unique news articles by the source countries is shown in Figure 1, and the distribution of news articles (including duplicates) over time is shown in Figure 2, which is intended to show the degree of excitement that accompanied the 2017 and 2019 BRI-associated forums in Beijing.

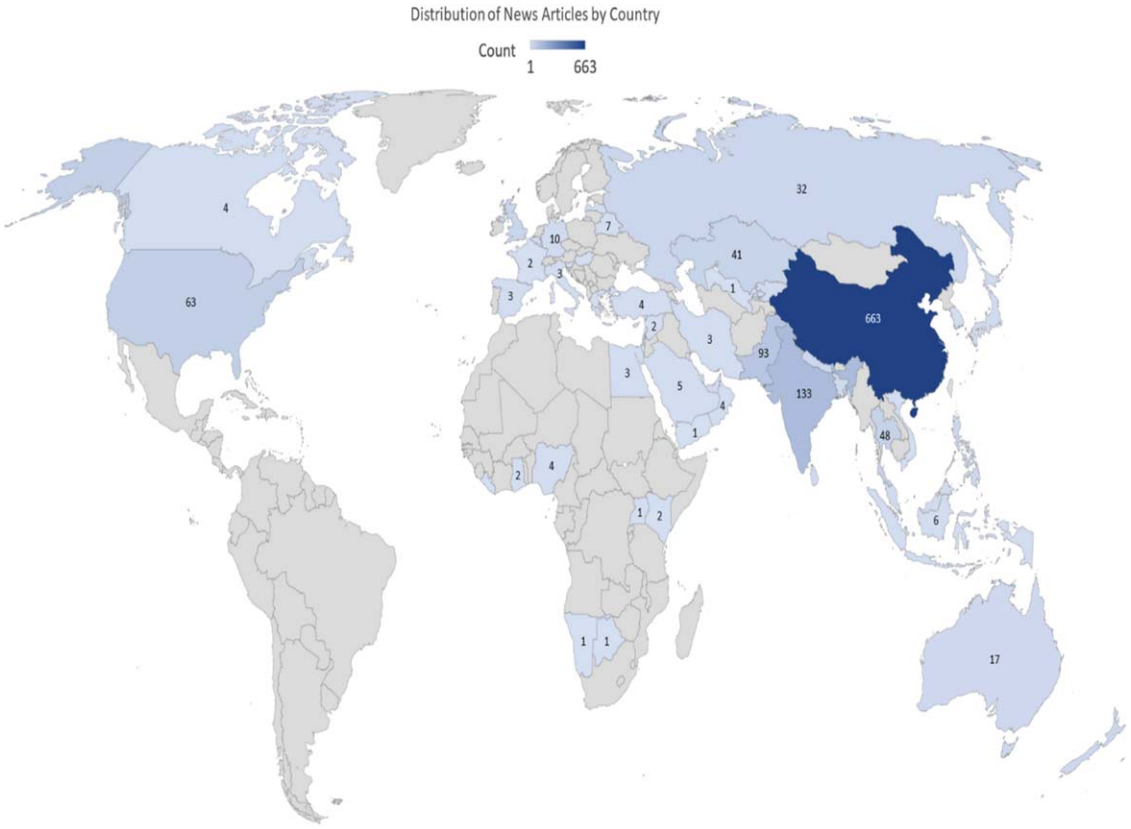

Figure 1: Distribution of news articles by country.

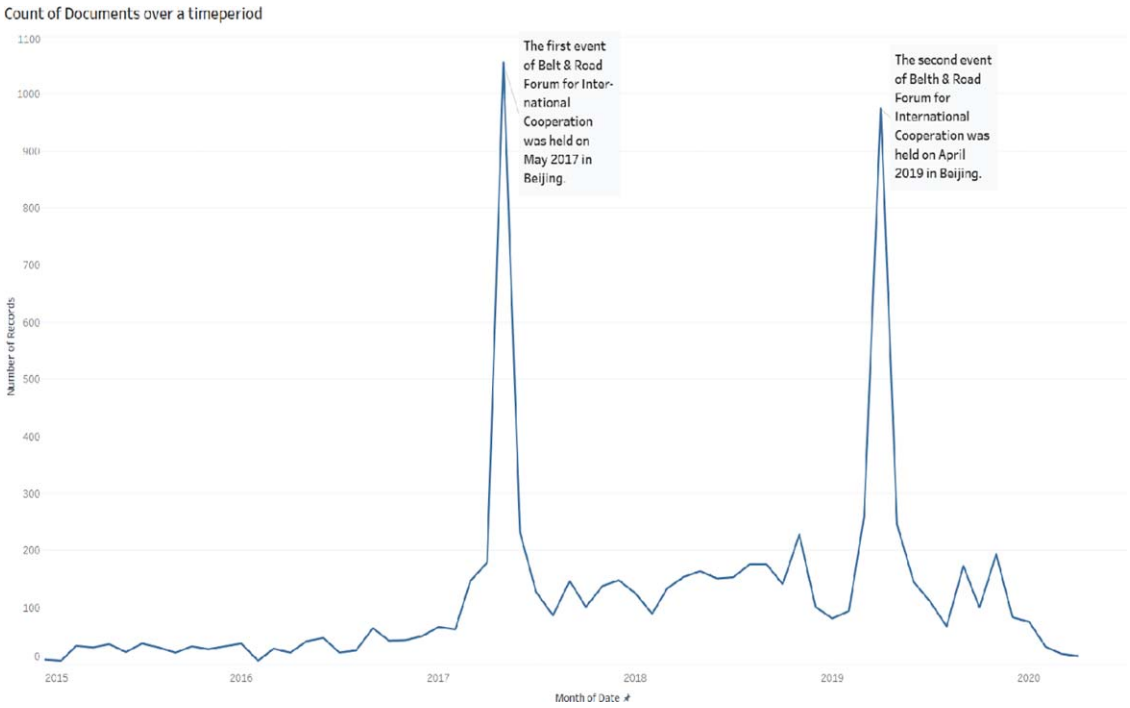

Figure 2: Distribution of the number of news articles over time.

We divided the countries into three regional categories: China, countries participating in BRI (henceforth referred to as participating countries), and countries not participating in BRI (henceforth referred to as nonparticipating countries). We felt that it was prudent to divide the countries into these three categories because China is the original initiator and sponsor of the project, whereas nonparticipating countries have maintained opposition to the project for various reasons, including geopolitics, and, finally, participating countries that have at least on paper signed on to this project. At the time of this report, 138 countries have signed cooperation documents with China for the BRI (Table 1).

These agreements are exploratory and have not necessarily translated into actual projects. It, therefore, is assumed that they are open to actual projects but are not essentially committed. China, as the sponsor, could be argued to have a relatively positive outlook of the project, although those countries that have consistently held an unfavorable view could be considered as being committed to an opposing perspective. Finally, those that have signed the agreements are probably the ones most open to different perspectives. This paper, therefore, separates the sentiments across the three groups to avoid cross bias among them. The countries that we categorized as participating and nonparticipating are shown in Figure 3.

5. Analysis

We used sentiment analysis to detect the emotions from the news articles to address our research questions. Sentiment analysis (also referred to as opinion mining) is a NLP technique that attempts to categorize the emotions and sentiments in a block of text. Most sentiment analysis tools will classify the sentiment as positive, negative, or neutral, and will also provide indexes for affective states, for example, anger, sadness, happiness. Sentiment analysis has been widely used to mine emotions from social media posts and news articles, and to effectively identify depression, anxiety, and other emotions (De Choudhury *et al.* 2013).

There are two main approaches to extracting sentiment from text. The lexicon-based approach uses a dictionary of words annotated with their sentiment polarities, whereas the text classification approach involves building classifiers from labeled instances of texts. Lexicon or dictionary-based approaches work well when there are insufficient human classified data or when human classification is time-consuming and expensive. The lexicon-based approach has several important advantages: first, once the dictionary is selected, researcher subjectivity is avoided; second, the method scales to large samples; and third, because the dictionaries are publicly available, it is easier to replicate the analysis of other researchers (Loughran and McDonald 2016).

Table 1: The list of countries that have signed cooperation documents for the Belt and Road Initiative (Belt and Road Portal 2020).

No.	Country	No.	Country	No.	Country	No.	Country
1	Afghanistan	36	El Salvador	71	Luxembourg	106	Senegal
2	Albania	37	Equatorial Guinea	72	Madagascar	107	Serbia
3	Algeria	38	Estonia	73	Malaysia	108	Seychelles
4	Angola	39	Ethiopia	74	Maldives	109	Sierra Leone
5	Antigua and Barbuda	40	Federated States of Micronesia	75	Mali	110	Singapore
6	Armenia	41	Fiji	76	Malta	111	Slovakia
7	Austria	42	Gabon	77	Mauritania	112	Slovenia
8	Azerbaijan	43	Gambia	78	Moldova	113	Solomon Islands
9	Bahrain	44	Georgia	79	Mongolia	114	Somalia
10	Bangladesh	45	Ghana	80	Montenegro	115	South Africa
11	Barbados	46	Greece	81	Morocco	116	South Sudan
12	Belarus	47	Grenada	82	Mozambique	117	Sri Lanka
13	Benin	48	Guinea	83	Myanmar	118	Sudan
14	BiH (Bosnia and Herzegovina)	49	Guyana	84	Namibia	119	Suriname
15	Bolivia	50	Hungary	85	Nepal	120	Tajikistan
16	Brunei	51	Indonesia	86	New Zealand	121	Tanzania
17	Bulgaria	52	Iran	87	Niger	122	Thailand
18	Burundi	53	Iraq	88	Nigeria	123	Togo
19	Cambodia	54	Israel	89	Niue	124	Tonga
20	Cameroon	55	Italy	90	North Macedonia	125	Trinidad Tobago
21	Cape Verde	56	Ivory Coast	91	Oman	126	Tunisia
22	Chad	57	Jamaica	92	Pakistan	127	Turkey
23	Chile	58	Kazakhstan	93	Panama	128	Uganda
24	Comoros	59	Kenya	94	Papua New Guinea	129	Ukraine
25	Costa Rica	60	Kiribati	95	Peru	130	United Arab Emirates
26	Croatia	61	Korea	96	Philippines	131	Uruguay
27	Cuba	62	Kuwait	97	Poland	132	Uzbekistan
28	Cyprus	63	Kyrgyzstan	98	Portugal	133	Vanuatu
29	Czech Republic	64	Laos	99	Qatar	134	Venezuela
30	Djibouti	65	Latvia	100	Republic of Congo	135	Vietnam
31	Dominic	66	Lebanon	101	Romania	136	Yemen
32	Dominica	67	Lesotho	102	Russia	137	Zambia
33	East Timor	68	Liberia	103	Rwanda	138	Zimbabwe
34	Ecuador	69	Libya	104	Samoa		
35	Egypt	70	Lithuania	105	Saudi Arabia		

We used the lexicon-based approach in this study due to its inherent advantages. It would be time-consuming and impractical to manually classify the sentiment in the news articles to create a large enough training dataset. Moreover, the manual classification approach would introduce human bias. There are several sentiment lexicons publicly available. We used a popular lexicon called the NRC lexicon (Mohammad and Turney 2013), which consists of 14,182 words, 2,317 positive, and 3,338 negative.

We preprocessed the data by removing stop words, punctuation, numbers, white spaces, and other words that would not convey sentiment. The sentiment analysis was done on the preprocessed data. The sentiment analysis output was a count of positive and negative sentiment words by month and by the three regional categories (China, participating countries, and nonparticipating countries). We computed a positive, negative, and overall sentiment index from this output by month and by regional category. We computed an index of positive sentiment by normalizing the number of positive sentiment words by the total number of words in that month. The negative sentiment index was computed in the same way as the positive sentiment index. An overall net sentiment index was computed by subtracting the count of negative sentiment words from the count of positive sentiment words and normalizing

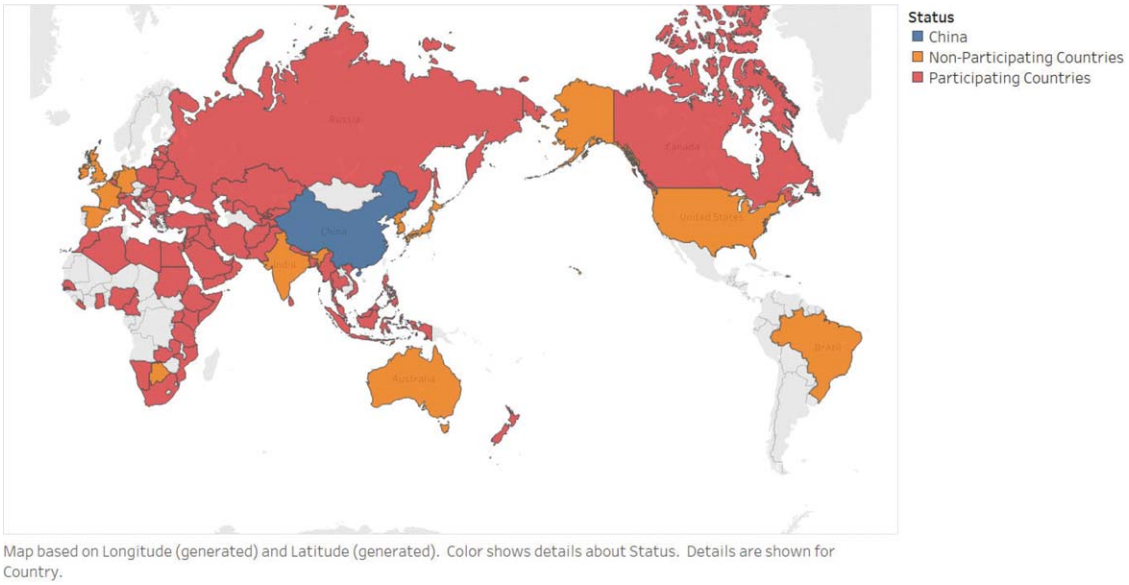

Figure 3: Participating and nonparticipating countries. Note: countries that we were not able to verify the status of have been grayed out.

this difference by the total number of words. We used the methodology described above to determine positive and negative sentiment of the BRI consistently across China, the participating countries, and the nonparticipating countries. We used the R software to perform the analysis.

6. Results and Discussion

The overall sentiment aggregated across all the countries that were part of our analysis is shown in Figure 4. As can be noted from a visual inspection of the graph, the overall positive sentiment is greater than the overall negative sentiment, thus, which leads to a net positive overall sentiment for BRI. The positive sentiment seems to be declining over time in Figures 4–7 that is, overall, China, participating countries, and nonparticipating countries, whereas the negative sentiment has held steady. The graph also shows peaks in overall positive sentiment during 2016.

Trend lines were fit to the data by using time series regressions to evaluate the trend in sentiments in a statistical manner. The dependent variables were the overall, positive, and negative sentiments, and the independent variable was time. The parameters of the time series regressions for the trend lines shown in Figures 4–7 are shown in Table 2.

The regression results show statistically significant ($p < 0.001$) negative slopes for overall sentiment and positive sentiment across all three groups: China, participating countries, and nonparticipating countries. The regression model p values were also statistically significant at the 99% level. The slopes for the negative sentiment for China and the nonparticipating countries were not statistically significant, which indicates no significant change in the negative sentiment. Thus, the statistical analysis corroborates in a more rigorous manner what can be gleaned by a visual inspection of the graphs. A further statistical test was conducted to see if the mean sentiments were different among the three groups: China, participating countries, and nonparticipating countries. The results of the analysis of variance for difference of means are shown in Table 3.

The analysis of variance results show that there was a statistically significant ($p < 0.001$) difference in all three means among the three groups. Post-hoc analysis was performed by using the Tukey HSD (honestly significant difference) multiple comparison tests. The Tukey tests showed a statistically significant ($p < 0.001$) difference in the mean sentiments among all three pairs (China versus participating countries, China versus nonparticipating countries, and participating versus nonparticipating countries).

One reason for the overall positive sentiment being greater than the overall negative sentiment could be that China has created an overall positive outlook on the program. This can be confirmed by the fact that in fewer than seven years since the project was initially proposed by President Xi Jin Ping of China, by some counts, there are 118 projects planned (Hielscher and Ibold 2020), spread across 138 countries (Belt and Road Portal 2020). However, these figures are difficult to verify because the projects are of varying types across multiple countries, and their status is not clear. There are some projects that have been agreed to on paper but have not progressed beyond the planning

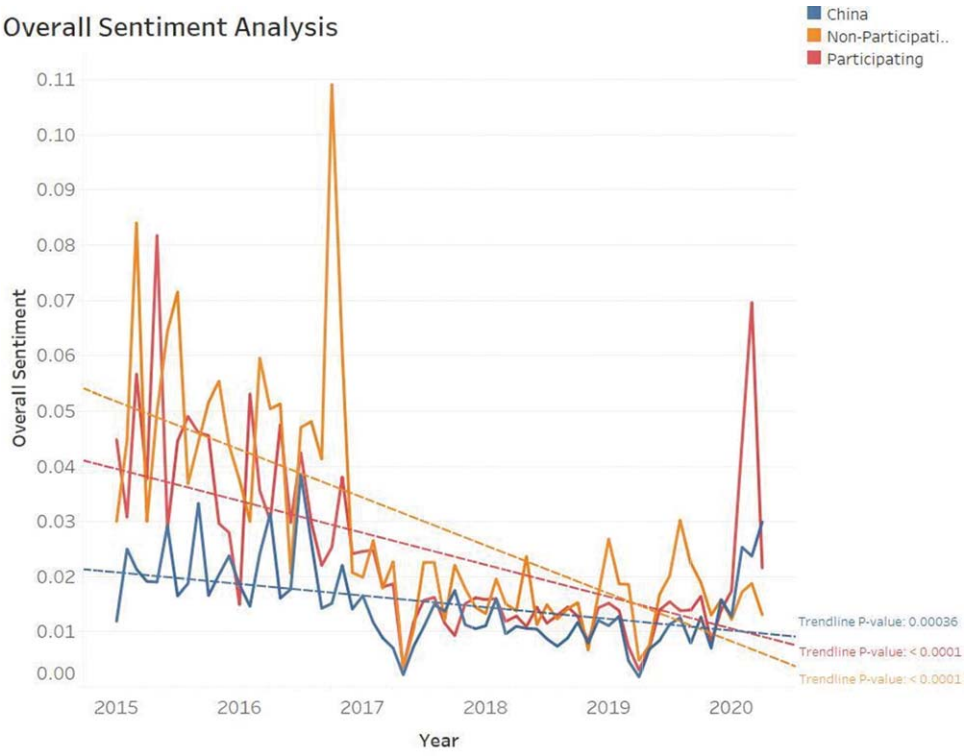

Figure 4: Overall sentiment across all countries.

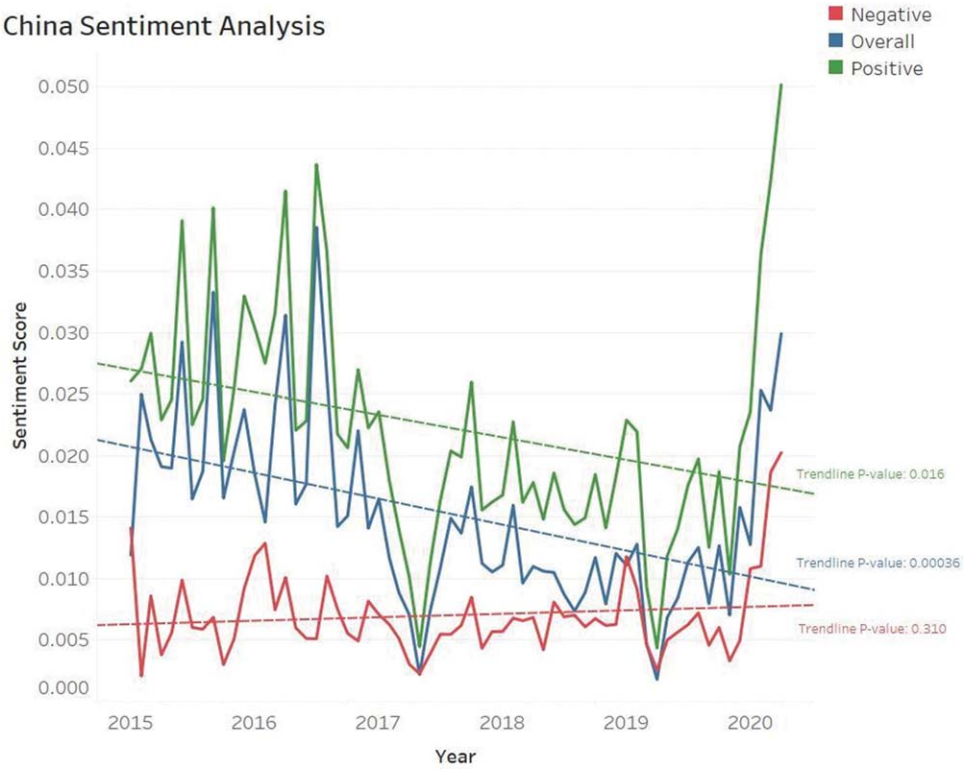

Figure 5: Sentiment analysis of news articles from China.

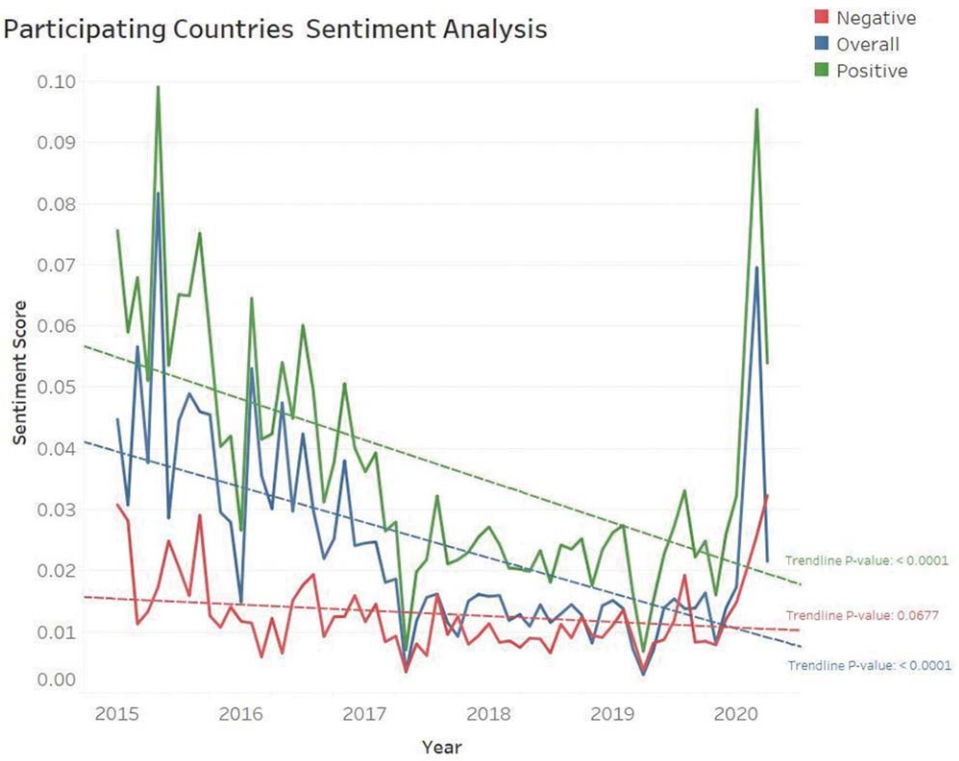

Figure 6: Sentiment analysis of news articles from participating countries.

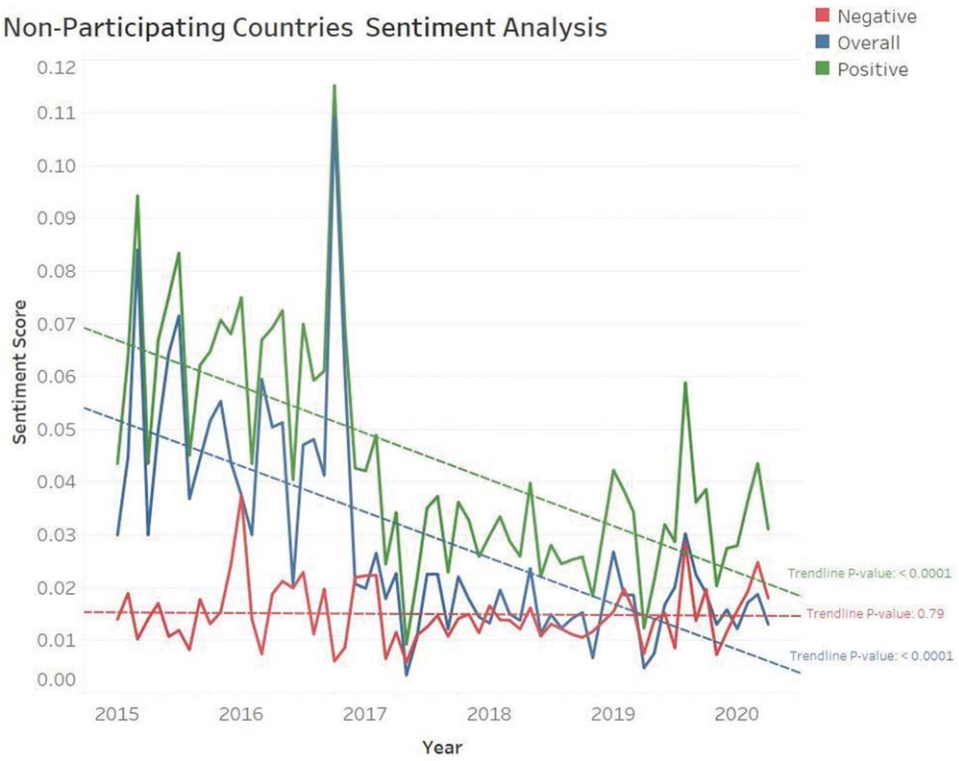

Figure 7: Sentiment analysis of news articles from nonparticipating countries.

Table 2: Time series regression (trend line) parameters.

Dependent Variable	Group	Slope	Intercept	R^2	Regression Model *p* Value
Overall sentiment	China	–0.00021 (6.37 e-05)***	0.0216	0.1544	0.0021***
	Participating	–0.00068 (7.78 e-05)***	0.0441	0.5725	4.6 e-12***
	Nonparticipating	–0.0007 (0.00014) ***	0.0519	0.3125	4.9 e-06***
Positive sentiment	China	–0.00023 (7.11 e-05)***	0.0288	0.1504	0.0024***
	Participating	–0.0009 (9.22 e-05)***	0.0615	0.5998	6.25 e-13***
	Nonparticipating	–0.00073 (0.00014)***	0.0677	0.3233	2.63 e-06***
Negative sentiment	China	−0.00002 (1.9 e-05)	0.0072	0.0193	0.2939
	Participating	–0.00017(3.75 e-05)***	0.0013	0.2721	2.26 e-05***
	Nonparticipating	–3.6 e-05 (4.37 e-05)	0.0158	0.012	0.41598

Standard errors are reported in parentheses.
***Indicates significance at the 99% level

Table 3: The analysis of variance test for comparison of mean sentiments.

	F Statistic	*p*
Overall sentiment	13.78	2.79 e-06***
Positive sentiment	26.5	8.98 e-11***
Negative sentiment	43.41	5.07 e-16***

***Indicates significance at the 99% level.

stage or have been scaled down (Chandran 2019). Other projects have been canceled and then restarted under different terms, for example, the Malaysian East Coast Rail Link (Fook 2019) and the Sri Lankan Hambantota Port Development Project (Patrick 2017). Finally, some projects have been credited to the BRI but may actually not be part of it (e.g., Mumbai Metro Line 4: India has consistently refused to join BRI). There is no denying that there is considerable excitement about the BRI program, and the more than 7,000 news articles that we collected are a testament to that. Below, we discuss some possible reasons why the overall positive sentiment seems to be declining over time, whereas the overall negative sentiment has held relatively steady over time.

One of the biggest criticisms of the BRI is the debt trap that it is accused of creating for the participating countries. The unique selling point of China as a financier was its no-strings-attached approach to lending money, unlike other financial organizations, for example, the World Bank. However, this has led to some unviable projects being undertaken and funded at very high interest rates (Wibisono 2019). The resultant debt has caused a backlash in many democratic countries with the freedom for stakeholders to voice their opinions (Balding 2018). These experiences seem to have contributed to the project's declining sentiment (Holland 2018; Rakhmat and Indramawan 2019).

Initial "euphoria" can best be witnessed by the highest positive sentiment in the eventually participating countries during May 2015. A review of the articles published during that period reveals the outreach by the Chinese government with regard to the project and the positive aspects highlighted by them. Another interesting result is that, in general, across all countries, the variation in positive sentiment (SD [Standard Deviation] = 0.017) is greater than the variation in negative sentiment (SD = 0.006). Variation in the overall sentiment seems to be highest among nonparticipating countries (SD = 0.023) as opposed to China (SD = 0.008) and the participating countries (SD = 0.015).

We have seen peaks in positive sentiment after the announcement of any new national agreement or project, which thus causes significant variation. The negative sentiment, however, seems to be steady across the period. For example, August 2019, which is a peak of both positive and negative sentiment for participating and nonparticipating countries, has announcements and updates of several national agreements (Nepal, Myanmar, Saudi Arabia, Mali, San Marino, Botswana, Thailand, Malaysia, Cambodia, Bulgaria, Turkey, Morocco, Pakistan, Iran, Kazakhstan, Nigeria, Russia, Uzbek, and the Philippines), along with the usual cautionary note, perhaps with a higher tone (e.g., Secretary Bolton's (Sputnik, 2019) and General David Petraeus' (The Australian, 2019) remarks). All three regional categories show a significant dip in positive and negative sentiment for May 2017 and April 2019. May 2017 and April 2019 were periods of Belt and Road Forums for International Cooperation. These periods had the highest number of articles (approximately 1000 in both instances), but they seemed to have a generally neutral and/or factual tone to the reporting.

6.1 Key Drivers of Sentiment

We now analyze and discuss the key drivers of the sentiments by investigating the prominent peaks and valleys in the sentiment across time. The following analysis uses word clouds to assist in identifying the key drivers of sentiments. These clouds highlight the words that drew the sentiment. The analysis follows the timeline with the corresponding word clouds. The following features are applied to the formation of the word clouds:

1. All words are processed in lower case
2. Common words, conjunctions, and symbols were ignored, e.g., and, is, "." (period),! (exclamation) etc.
3. More frequently mentioned words have a larger font.
4. Similar frequency words have the same color

China February 2016

6.1.1 China, February 2016

February 2016 was when both positive and negative (mixed) sentiments peaked in the Chinese media. These sentiments highlighted the competition faced by the Chinese in the global railroad infrastructure markets, particularly from the long-established Japanese. The positive sentiment seemed to have originated from its success in outbidding the Japanese in the contract to construct a high-speed railway linking Jakarta and Bandung in Indonesia. This positive event from the Chinese perspective seemed to have been balanced by the news of the Japanese succeeding in its bid to build a high-speed line between Mumbai and Ahmedabad in India. The success of the Japanese was seen as a threat from a strategic perspective by the Chinese.

Non-Participating Countries October 2016

6.1.2 Nonparticipating Countries, October 2016

October 2016 saw a spike in positive sentiment in articles from nonparticipating countries. This upsurge in sentiment was due to limited extreme data (highly opinionated articles) that month that skewed the results.

Participating Countries **Non-Participating Countries**

December 2016

6.1.3 Participating Countries and Nonparticipating Countries, December 2016

December 2016 saw increased negative sentiments in both participating and nonparticipating countries. This negativity had to do with the regional conflict between India (nonparticipating) and Pakistan (participating), and the implications of the BRI program and its regional component, the China-Pakistan Economic Corridor project on the two countries. The possible joining of Bangladesh to this program was also reported with concern by the Indian side.

Participating Countries August 2017

6.1.4 Participating Countries, August 2017

During August 2017, there was a peak in negative sentiment from the participating countries. The words "health," "thailand," and "minister" appear in this word cloud. China hosted the belt and road high-level meeting for health cooperation in Beijing, attended by ministers from 30 different countries, along with 300 health officials and World

Health Organization leaders. However, three key ASEAN (Association of Southeast Asian Nations) leaders were not invited to the summit: the prime ministers of Thailand, Singapore, and Brunei. China's choice not to invite the aforementioned prime ministers is the source of much negative sentiment from participating countries. Many believe the decision to omit the ASEAN leaders was in response to delays in the Thai-Chinese high-speed railway project despite the two nations conducting 18 rounds of (unsuccessful) negotiations.

China October 2017

6.1.5 China, October 2017

October 2017 illustrates a spike in positive sentiment from Chinese sources. China hosted a two-day "EU-Eurasia-China business summit" that focused on promoting the growth of the BRI into eastern areas where it had not yet started construction. During the summit, the Greek president Prokopis Pavlopoulos stated his support for becoming economically involved in the BRI. Positive sentiment in this month was primarily due to the potential growth of the BRI as a result of the successful summit.

China January 2019

6.1.6 China, January 2019

There was a peak in negative sentiment in January 2019. A Chinese article published on the seventh shares the results of a survey of respondents from ASEAN countries (Kim et al., 2019), where 70 percent of the participants agreed with the statement, "Southeast Asian countries should be cautious in negotiating with China on its flagship Belt and Road Initiative to avoid being trapped in unsustainable debt." A Chinese news article (Wilson, 2019) also expressed negative sentiment in response to Australia's reluctance to sign BRI agreements, despite their neighbor New Zealand's good relationship with China.

Non-Participating February 2019

6.1.7 Nonparticipating Countries, February 2019

February 2019 saw a peak in negative sentiment from nonparticipating and participating countries. The word cloud features "security," "terrorist," and "taliban." The word cloud and peaks in negative sentiment are in response to an attack from terrorist Masood Azhar-led Jaish-e-Mohammed (Jem) on an Indian CRPF (Central Reserve Police Force) convoy at Pulwama. The attack killed 40 people and injured many more. Many articles from participating and nonparticipating sources in August focused on the rising tensions with China and discussed the potential negative impacts of the BRI.

Non-Participating Countries August 2019

6.1.8 Nonparticipating Countries, August 2019

In August 2019, there was a peak in mixed sentiment from nonparticipating nations and a negative sentiment from participating countries. Australian sources from this time expressed heavy negative sentiment in response to the choice of New Zealand to participate in the BRI. An Indian article discussed the upcoming "BRI, China Pakistan Economic Corridor, and TransRegional Integration" conference in Pakistan with positive sentiment and explained how India could contribute to Asia's success as a continent if it participates in the BRI. While acknowledging the potential benefits that BRI could bring in, an article from Myanmar (Thein, 2019) advocated caution to ensure that the local population benefited from the projects.

6.1.9 China, Participating Countries, and Nonparticipating Countries, March 2020

During March 2020, there was a peak in both types of sentiment from all the groups. Many articles from China and their allies discussed how the COVID-19 pandemic will affect BRI's international progress. Many of these articles share the positive sentiment that, because COVID-19 may have a temporary impact on the BRI's performance, the setback will be minor. Some articles (Jianguo, 2020; Zheng & Lo, 2020) even discuss how China demonstrated good emergency preparedness and call the chain of events a success.

The sentiments around BRI could well be considered a proxy for the general perception of Chinese foreign policy global endeavors. China has increasingly been asserting its intent to take over the leadership mantle from the United States (Disis and He 2021; The Policy Planning Staff 2020). The BRI program has both economic as well as political consequences for the countries involved as well as those not involved in it. In China's own neighborhood, BRI-supporting countries, such as Indonesia, Malaysia, Myanmar, Pakistan, Philippines, Sri Lanka, Thailand, and Vietnam, have diverse sentiments about the viability and implications of the program (Chao 2021; Chin 2021; Cox *et al.* 2018; Moramudali 2016; Mursitama and Ying 2021; Punyaratabandhu and Swaspitchayaskun 2021; Qianqian and Yijun 2021; Soong and Aung 2021; Vu, Soong, and Nguyen 2021). Because the implications of this program go beyond transportation efficiency, stakeholder perceptions of this program have takeaways for global entities (commercial and political) in the way that they strategize their future. The size of the program and the constituent national projects have created financial dependence on China in many of the vulnerable economies.

The much-discussed and analyzed case of the Sri Lankan port of Hambantota, which ended up being leased to China for ninety-nine years in lieu of debt, is an example of this (Carrai 2018). If the trend of increasing economic and the resultant strategic ties between China and BRI host countries continues, it will undoubtedly have implications for the United States and other countries in the region that have uneasy relations with China, such as Australia, India, and Japan. These concerns are likely to be shared with other countries beyond Asia, in Europe and Africa. With BRI, China is also entering the Balkan region (Vangeli 2020), with implications for countries in that region. This entry leads to a three-way struggle for influence among Europe, Russia, and China, with the latter two developing a closer relation. The scope of BRI goes beyond infrastructure into the digital domain (Digital BRI) with the push by China for its 5 G telecommunication technologies and hardware (Bartholomew 2020). The influence gained by China can help it open new markets for its 5 G equipment both today as well as tomorrow.

COVID-19 has caused some slowdown in the BRI projects that may well be temporary because practically all the affected nations expect the pandemic to eventually subside (although not clear when). In fact, the pandemic itself

has created an opportunity for China to conduct health diplomacy in the form of a Health Silk Road (Chow-Bing 2020). China has been the go-to source for critical personal protective equipment and a likely source for vaccines. This explains the positive sentiment with regard to BRI, with "cooperation" being the key term for China and the participating countries in March 2020.

7. Conclusion

The results of our study can help in policy decisions. Countries can use the results of our study and adapt our analysis to understand how their public is viewing the BRI initiative and use this information in making policy decisions. Our study is among the first to use sentiment analysis, a NPL technique, to explore the global perception of the BRI. Our findings of a declining and increasingly volatile positive sentiments and a steady and/or stable negative sentiment sheds light on how the global perspective on BRI has changed over time. The spike in interest around the period when BRI-related forums were hosted in Beijing, China, indicates the efficacy of those initiatives by the Chinese government. However, the declining positive sentiment seems to suggest a more cautious perspective that could be the result of the outcomes of the current projects or the geopolitical environment. The fact that the sentiment is still overall positive bears testament to the fact that countries appreciate this initiative of President Xi. This positive sentiment is a takeaway for what could be a successful foreign policy.

The US withdrawal from Afghanistan has opened the doors for China's entry offering support for the war-torn country in the form of infrastructure development under the BRI. Afghanistan is rich in mineral wealth and can provide a connection to central Asia from the port of Gwadar in Pakistan, which has been developed as a part of the BRI and the China-Pakistan Economic Corridor. China has stepped up as one of the first countries to develop relations with the Taliban government in Afghanistan. These developments will no doubt have implications for the BRI program. Those committed to either supporting or rejecting the BRI-affiliated projects will probably continue to do so. However, it would be interesting to see the sentiments emerging from countries that were open to but did not actually initiate any projects. Our research makes it possible for a repeat of this analysis by using the latest news articles to show the change if any in attitudes to the BRI as well as to China.

One of our study's limitations is that the sentiment analysis does not lead to causal inferences. A further limitation is that our results are dependent on the lexicon we picked for the sentiment analysis, although the lexicon we picked is a popular and comprehensive one. Further research could be performed with different lexicons to assess if the results change significantly with the choice of the lexicon. Although our in-depth analysis of the drivers of the sentiment sheds light on the causes, further research needs to be performed to explore in detail the causes for the changes in sentiments over time. Another limitation is that we have only considered English language news articles in our analysis. An analysis of local language news articles is currently infeasible with the current technology because lexicons for local languages are not available. But, this could be an avenue for future research as the technology improves and local lexicons become available. Other avenues for further research include drilling down to the individual country level and exploring changes in sentiment and further exploring the change in sentiment around certain major BRI-related events, particularly project cancellations.

References

Adrien, M. 2019. "OFNRS - Observer, Analyser et Conseiller." Accessed January 23, 2021. https://observatoirenrs.com/.

Ali, Z., Ö. Gökçe, M. Binark, and A. D. Gidreta. 2020. "China-Pakistan Economic Corridor and Technicians of Opinion in Pakistani Twittersphere: A Thematic Content Analysis." *Asya Araştırmaları Uluslararası Sosyal Bilimler Dergisi* **4**, no. 1: 9–28.

Arifon, O., Z. A. Huang, Y. Zheng, and A. Zyw Melo. 2019. "Comparing Chinese and European Discourses Regarding the 'Belt and Road Initiative'." *Revue Française Des Sciences de L'information et de la Communication* **17**.

Balding, C. 2018. "Why Democracies Are Turning against Belt and Road." *Foreign Affairs* **97**, no. 5.

Bartholomew, C. 2020. "China and 5G." *Issues in Science and Technology* **36**, no. 2: 50–57. Accessed January 23, 2021. https://www.proquest.
com/scholarly-journals/china-5g/docview/2452125915/se-2.

Belt and Road Portal. 2020. "List of countries that have signed cooperation documents with China for the Belt and Road Initiative." Accessed January 23, 2021. https://www.yidaiyilu.gov.cn/info/iList.jsp?tm_id=126&cat_id=10122&info_id=77298

Bērziņa-Čerenkova, U. A. 2016. "BRI instead of OBOR – China edits the English Name of its most ambitious international project." Accessed January 23, 2021. http://liia.lv/en/analysis/bri-instead-of-obor-china-edits-the-english-name-of-its-most-ambitious-international-project-532

Bloomberg. 2019. "Analysis |China's Belt and Road is getting a reboot. Here's why." *Bloomberg News*. Accessed https://www.washingtonpost.com/business/chinas-belt-and-road-is-getting-a-reboot-heres-why/2019/08/14/6725fe4e-be63-11e9-a8b0-7ed8a0d5dc5d_story.html

Carrai, M. A. 2018. "China's Malleable Sovereignty Along the Belt and Road Initiative: The Case of the 99-Year Chinese Lease of Hambantota Port." *New York University Journal of International Law and Politics* **51**: 1061.

Chandra, J. K., E. Cambria, and A. Nanetti. 2020. "One Belt, One Road, One Sentiment? A Hybrid Approach to Gauging Public Opinions on the New Silk Road Initiative." Paper Presented at the 2020 International Conference on Data Mining Workshops (ICDMW), Sorrento, Italy, November 17–20.

Chandran, N. 2019. "Fears of excessive debt drive more countries to cut down their Belt and Road investments." Accessed January 23, 2021. https://www.cnbc.com/2019/01/18/countries-are-reducing-belt-and-road-investments-over-financing-fears.html.

Chao, W.-C. 2021. "The Philippines' Perception and Strategy for China's Belt and Road Initiative Expansion: Hedging with Balancing." *The Chinese Economy* **54**, no. 1: 48. doi: 10.1080/10971475.2020.1809817

Chatzky, A., and J. McBride. 2019. 05/21/2019). "China's massive Belt and Road initiative." Accessed January 23, 2021. https://www.cfr.org/backgrounder/chinas-massive-belt-and-road-initiative.

Chin, K. F. 2021. "Malaysia's Perception and Strategy toward China's BRI Expansion: Continuity or Change?" *The Chinese Economy* **54**, no. 1: 9–11. doi: 10.1080/10971475.2020.1809814

Chow-Bing, N. 2020. "*COVID-19, Belt and Road Initiative and the Health Silk Road: Implications for Southeast Asia*." Paper Presented at the 4th NACAI International Symposium 21 November 2020 University of Yangon (Centennial).

Cox, M., T. S. M. Majid, Y. Jie, J. Yan, and H. Hamzah. 2018. "China's Belt and Road Initiative (BRI) and *Southeast* Asia." *CIMB ASEAN Research Institute* **47**.

Das, S., G. Medina, L. Minjares-Kyle, and Z. Elgart. 2018. "Social Media Hashtags Associated with Bike Commuting: Applying Natural Language Processing Tools (No. 18-03545)." In *Paper Presented at the Transportation Research Board 97th Annual Meeting*, Washington DC, January 7–11, 2018.

De Choudhury, M., M. Gamon, S. Counts, and E. Horvitz. 2013. "Predicting Depression via Social Media." In *Proceedings of the International AAAI Conference on Web and Social Media* **7**, no. 1: 128–137.

Disis, J., and L. He. 2021. "China is rehearsing for when it overtakes America." *CNN Business*. Accessed January 23, 2021. https://www.cnn.com/2021/01/26/economy/china-xi-economy-intl-hnk/index.html.

Evans-Cowley, J. S., and G. Griffin. 2012. "Microparticipation with Social Media for Community Engagement in Transportation Planning." *Transportation Research Record: Journal of the Transportation Research Board* **2307**, no. 1: 90–98. doi: 10.3141/2307-10

Fook, L. L. 2019. China-Malaysia relations back on track? Accessed January 23, 2021. http://hdl.handle.net/11540/10256.

Griffiths, R. T. 2017. *Revitalising the Silk Road: China's Belt and Road Initiative*. Leiden, The Netherlands: Hipe Publications.

Gu, Y., Z. S. Qian, and F. Chen. 2016. "From Twitter to Detector: Real-Time Traffic Incident Detection Using Social Media Data." *Transportation Research Part C: emerging Technologies* **67**: 321–42. doi: 10.1016/j.trc.2016.02.011

Hielscher, L., and S. Ibold. 2020. "Belt and Road initiative." Accessed January 23, 2021. https://www.beltroad-initiative.com/projects/

Holland, T. 2018. "Why borrowers on China's Belt and Road will go from Euphoria to depression." *South China Morning Post*. Accessed https://www.scmp.com/week-asia/opinion/article/2138499/why-borrowers-chinas-belt-and-road-will-go-euphoria-depression.

Hurley, J., S. Morris, and G. Portelance. 2019. "Examining the Debt Implications of the Belt and Road Initiative from a Policy Perspective." *Journal of Infrastructure, Policy and Development* **3**, no. 1: 139–75. doi: 10.24294/jipd.v3i1.1123

Indermit Gill, S. V. L., and L. Mathilde. 2019. "Winners and losers along China's Belt and Road." Accessed January 23, 2021. https://www.brookings.edu/blog/future-development/2019/06/21/winners-and-losers-along-chinas-belt-and-road/

Jianguo, W. (2020, 03/06/2020). Coronavirus an opportunity for BRI healthcare cooperation. Global Times. Accessed through Factiva 04/23/2020.

Kim, J., S. Lewis, and C. Fernandez. (2019, 01/07/2019). Southeast Asia wary of China's Belt and Road project, sceptical of U.S.: survey. Macau Business Daily. Accessed through Factiva 03/11/2019.

Li, J., Y. Yao, Y. Xu, J. Li, L. Wei, and X. Zhu. 2019. "Consumer's Risk Perception on the Belt and Road Countries: Evidence from the Cross-Border e-Commerce." *Electronic Commerce Research* **19**, no. 4: 823–40. doi: 10.1007/s10660-019-09342-x

Li, Y., Y. Zhang, L. A. Tiffany, R. Chen, M. Cai, and J. Liu. 2021. "Synthesizing Social and Environmental Sensing to Monitor the Impact of Large-Scale Infrastructure Development." *Environmental Science and Policy* **124**: 527–40. https://doi.org/10.1016/j.envsci.2021.07.020. doi: 10.1016/j.envsci.2021.07.020

Liu, Y., Y. Li, and W. Li. 2019. "Natural Language Processing Approach for Appraisal of Passenger Satisfaction and Service Quality of Public Transportation." *IET Intelligent Transport Systems* **13**, no. 11: 1701–1707. doi: 10.1049/iet-its.2019.0054

Loughran, T., and B. McDonald. 2016. "Textual Analysis in Accounting and Finance: A Survey." *Journal of Accounting Research* **54**, no. 4: 1187–230. doi: 10.1111/1475-679X.12123

Malik, T. H. 2020. "The Belt and Road Initiative (BRI) Project Legitimisation: The Rhetor's Innovation and the US Response." *Asian Journal of Comparative Politics* doi: 10.1177/2057891120959476

Mark, J. J. 2019. "Silk Road." Accessed January 23, 2021. https://www.ancient.eu/Silk_Road/

Mehrotra, S., and S. Roberts. 2018. "Identification and Validation of Themes from Vehicle Owner Complaints and Fatality Reports Using Text Analysis." In 97th Annual Meeting of the Transportation Research Board, Washington, DC.

Mohammad, S. M., and P. D. Turney. 2013. "Crowdsourcing a Word–Emotion Association Lexicon." *Computational Intelligence* **29**, no. 3: 436–465. doi: 10.1111/j.1467-8640.2012.00460.x

Moramudali, U. 2016. "Sri Lanka's Debt and China's money." *The Diplomat.* Accessed January 23, 2021. https://bandapost.org/wp-content/uploads/pdf/Dec%20Srilanka%20debt%20and%20China%20money.pdf.

Mursitama, T. N., and Y. Ying. 2021. "Indonesia's Perception and Strategy toward China's OBOR." *The Chinese Economy* **54**, no. 1: 35–47. doi: 10.1080/10971475.2020.1809816

Napitupulu, A., M.A. Embi, and B. Briando. 2020. "Public Sentiment Analysis on the Existence of Foreign Worker during the Covid 19 Pandemic." Paper Presented at the International Conference on Law and Human Rights. Virtual Conference, October 26–27, 2020.

Niu, W., and L. Wu. 2019. "Sentiment Analysis and Contrastive Experiments of Long News Texts." Paper Presented at the 2019 IEEE 4th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), Chengdu, China December 20–22.

Patrick, A. 2017. China-Sri Lanka Strategic Hambantota Port Deal. National Maritime Foundation, Accessed January 23, 2021. http://www.maritimeindia.org/View%20Profile/636276610966827339.pdf

Punyaratabandhu, P., and J. Swaspitchayaskun. 2021. "Thailand's Perception and Strategy toward China's BRI Expansion: Hedging with Cooperating." *The Chinese Economy* **54**, no. 1: 69–69. doi: 10.1080/10971475.2020.1809819

Qianqian, L., and L. Yijun. 2020. "The China-Pakistan Economic Corridor: The Pakistani Media Attitudes Perspective." *Technology in Society* **62**: 101303. https://doi.org/10.1016/j.techsoc.2020.101303. doi: 10.1016/j.techsoc.2020.101303

Rakhmat, M. Z., and D. Indramawan. 2019. Belt and Road Initiative: Challenging South and Southeast Asia – Analysis. Accessed January 23, 2021. https://www.eurasiareview.com/15112019-belt-and-road-initiative-challenging-south-and-south-east-asia-analysis/

Rithmire, M., and Y. Li. 2019. "Chinese Infrastructure Investments in Sri Lanka: A Pearl or a Teardrop on the Belt and Road?" Harvard Business School Case 719-046, January 2019. Accessed January 23, 2021. https://www.hbs.edu/faculty/Pages/item.aspx?num=55410

Soong, J.-J., and K. H. Aung. 2021. "Myanmar's Perception and Strategy toward China's BRI Expansion on Three Major Projects Development: Hedging Strategic Framework with State-Market-Society Analysis." *The Chinese Economy* **54**, no. 1: 20–34. doi: 10.1080/10971475.2020.1809815

Sputnik. (2019). Beijing Opposes Bolton's Remarks on Dangers of Chinese Investments - Foreign Ministry. Sputnik News Service.

Teo, H. C., A. Campos-Arceiz, B. V. Li, M. Wu, and A. M. Lechner. 2020. "Building a Green Belt and Road: A Systematic Review and Comparative Assessment of the Chinese and English-Language Literature." *Plos ONE* **15**, no. 9: e0239009. doi: 10.1371/journal.pone.0239009

The Australian. (2019, 08/14/2019). Petraeus crystallises challenges. The Australian. https://www.theaustralian.com.au/

The Policy Planning Staff. 2020. "The elements of the China Challenge." Accessed January 23, 2021. https://www.state.gov/wp-content/uploads/2020/11/20-02832-Elements-of-China-Challenge-508.pdf.

Thein, A. Z. P. (2019, 08/02/2019). The Belt and Road Initiative in Myanmar: staring down the dragon. Frontier Myanmar. Accessed through Factiva 03/11/2019.

UNESCO. 2019. "About the Silk Road | SILK ROADS." Accessed https://en.unesco.org/silkroad/about-silk-road.

Vangeli, A. 2020. "China's Belt and Road in the Balkans in the Post-COVID-19 Era." European Institute of the Mediterranean. *Mediterranean Yearbook 2020.* Accessed January 23, 2021. https://www.iemed.org/observatori/arees-danalisi/arxius-adjunts/anuari/med.2020/China_Belt_Road_Balkans_Anastas_Vangelis_IEMed_YearBook2020.pd.

Vu, V.-H., J.-J. Soong, and K.-N. Nguyen. 2021. "Vietnam's Perceptions and Strategies toward China's Belt and Road Initiative Expansion: Hedging with Resisting." *The Chinese Economy* **54**, no. 1: 56–68. doi: 10.1080/10971475.2020.1809818

Whitfield, S. 2007. "Was There a Silk Road?" *Asian Medicine* **3**, no. 2: 201–13. https://doi.org/10.1163/157342008X307839. doi: 10.1163/157342008X307839

Wibisono, A. N. 2019. "China's 'Belt and Road Initiative'." *In Sri Lanka: Debt Diplomacy in Hambantota Port Investment. Mandala: Jurnal Ilmu Hubungan Internasional* **2**, no. 2: 222–245.

Wilson, K. (2019, 01/24/2019). Canberra clings to outdated alliance. China Daily-Global Edition. Accessed through Factiva 01/23/2021.

Zheng, S., and K. Lo. (2020, 03/05/2020). Coronavirus: China keen to promote its success in controlling epidemic. South China Morning Post. Accessed through Factiva 04/23/2020.

Zhou, J., J. Wang, Q. Li, Y. Wen, Y. Zhang, and P. Lin. 2021. "A Public Opinion Analysis System Based on Emotion Analysis in Linyi." *Journal of Physics: Conference Series* **1757**, no. 1: 012116. doi: 10.1088/1742-6596/1757/1/012116