Voices from Venezuela: Examining Blogs to Study the Socio-Politico-Economic Crisis and Consequent Emigration

Esther Ledelle Mead, Muhammad Nihal Hussain, Mohammad Nooman, Samer Al-khateeb, Nitin Agarwal[†]
[†]Jerry L. Maulden-Entergy Chair Professor of Information Science
University of Arkansas at Little Rock, Little Rock, United States
{elmead, mnhussain, msnooman, sxalkhateeb, nxagarwal}@ualr.edu

Abstract—The objective of this research is to continue our journey into determining whether the blogosphere, as a type of social media platform, can be used to disseminate information regarding the socio-political views and concerns of citizens within a specific community. We expand upon our example case of focusing on information relative to the Venezuelan community regarding the current Venezuelan socio-economic crisis. Are Venezuelan blogs being used to discuss socio-political events and the quality of life concerns that are associated with the economic crisis in that community? Are Venezuelans using blogs to discuss possible migration away from the region as a result of these concerns? The Blogtrackers tool was used to analyze almost 30,000 Venezuelan blog posts collected between August 2003 and March 2017. Our analysis showed that the blogosphere is indeed being used as a platform by citizens to discuss these issues. We show how the posting frequency, sentiment, and keyword trends have changed over time relative to the changes in the socio-political landscape of the region and the events surrounding them. Of particular interest is the keyword trend analysis that shows that blogs are being used to discuss issues associated with quality of life factors and interest in migration away from Venezuela as a result of the crisis. We believe that this study can be used as a starting point to show the value of analyzing blogs in helping to identify humanitarian needs, in facilitating assistance efforts, and in drafting policy decisions.

Keywords- blogosphere; situation awareness; socio-political; social media data analysis; migration; Venezuela; Blogtrackers.

I. INTRODUCTION

This paper is an expanded version of the earlier work by Mead et al. [1], which established the basis for using blog analysis for studying socio-political awareness. Social media platforms, such as Twitter, Facebook, YouTube, and blogs have changed the way citizens express and share their sentiments regarding socio-political situations within their communities, and have created a space for citizen journalism [2]. These new mediums of communication have also allowed for citizen sentiment to be channelized from the online forum to the streets in the form of public debates and protests. Tufekci and Wilson [3], for example, demonstrated that participation in protests, both before and on the first day of the Tahrir Square demonstrations was elicited by information that citizens posted on blogs, Facebook, and Twitter.

A blog is a "personalized webpage, kept by the author in reverse chronological diary form" [4]. Blogs give people a social identity and are a medium for association, self-expression, and dissemination of information [5]. Blogs can be a very effective way to gain an in-depth understanding of issues and events due to the fact that recording, revisiting, and reflecting on the past is possible through previous blog entries. The ability for readers to leave comments in an interactive environment is an important part of blogging.

Blogs serve as an interactive platform for information exchange and discussion, and provide useful information about events [6]. Blogs serve as a way for citizens to gain situational awareness of the socio-political landscape of their environment, and data experts who track and analyze blogs can gain an understanding of the perspectives and intentions that exist among citizens. The Blogtrackers tool [7] was used to analyze a large dataset of Venezuelan blogs to determine whether the blogosphere is being used to disseminate information about issues stemming from the Venezuelan socio-economic crisis, and, if so, how this content is changing over time. This information is particularly helpful for emergency responders, and policy and decision makers leading humanitarian assistance efforts.

The rest of this paper is organized as follows. Section II reviews the currently published literature and related work regarding citizen use of social media platforms relative to the attainment of situational awareness of socio-political issues. Section III explains the methodology used such as the data collection process and the Blogtrackers tool (http://blogtrackers.host.ualr.edu/Blogtrackers/). Section IV provides a discussion of the data analysis results. Section V concludes the study outlining future research directions.

II. RELATED WORK

In this section we discuss the various previously published work related to the analysis of social media data as a means for studying socio-political awareness.

A. Informational Power of Social Media Platforms

Many sources highlight the power of social media platforms such as blogs, Twitter, Facebook and YouTube to be effective tools for allowing citizens to engage in socio-political scenarios such as obtaining information, disseminating information, participating in socio-political discussions, and becoming mobilized to act or to participate in impactful events [8, 9, 10, 11, 12, 13, 14, 15, 16]. Additionally, some sources focus on or add that the effectiveness of social media platforms as informational and motivational tools can be leveraged by organizations for crisis management and emergency response [8, 11, 15, 17, 18, 19, 20, 21]. Much of the literature highlights the power

of social media platforms by using the example of the 2010 protests in Egypt, followed by an Egyptian Revolution in 2011, wherein the public used social media to communicate their dissatisfaction with socio-political issues such as poverty, unemployment, corruption, high prices, repression, and human rights abuse [9, 10, 11, 16, 22, 23].

B. New Media Versus Traditional Media

Many sources refer to social media platforms as "New Media" [13, 16, 24], "Participatory Media" [9, 25], and "Modern Information and Communication Technologies (ICTs)" [13, 16, 22, 23], and draw a contrast with "Traditional Media" platforms such as television, newspaper, and radio (especially, state-controlled traditional media) [9, 11, 13, 14, 16, 22, 24, 25]. Valentini et al. [14] showed that crises tend to be communicated or framed differently within the writings of new media (specifically blogs) than they are within the writings of traditional media. Some sources contend that citizens are increasingly using social media sites for obtaining and disseminating socio-political information due to the perceived or real incompetence of or censoring by state-owned media outlets [16, 22, 24].

C. A Focus on the Blogosphere

Using the possibility of migration due to the Venezuelan economic crisis as a case study, Mead et al. established the basis for using blog analysis for studying socio-political awareness [1]. Before Mead et al., only a small pool of sources focus specifically on blogs as a means for obtaining and disseminating socio-political information [22, 26, 13, 14]. The number of blogs has been said to double every five months; they are easily found and can be accessed freely by anyone with an internet connection [13]. Al-Ani et al. [22] highlighted the power blogs played in mobilizing citizens during the Egyptian revolution of early 2011. Blogs represent a "counter-narrative" to the government-controlled media especially during times of crisis, and provide a means to voice dissent and to challenge authoritative power [22]. Blogs provide a means to potentially develop and maintain a strong sense of community among citizens interested in certain themed topics [26]. Additionally, blogs often-times represent "citizen-based news sources" that challenge traditional media in terms of the ability to form public opinion [14]. The information posted on blogs and the commentary reactions to the posts, therefore, have become of increasing interest to social media researchers.

D. Cybersecurity Issues

Some sources highlight the potential cybersecurity issues related to the use of social media platforms for spreading situational awareness in terms of socio-political issues [27, 25, 11]. Fearn [27] warns that "cybercriminals" are increasingly using programs called "bots" to attack social media users via Twitter, Facebook, and YouTube with negative comments and to spread disinformation or fake news via these platforms. Since the blogosphere is also a social media platform, the potential for blogs to be used nefariously by bots can be argued [25]. Goolsby et al. [25] also argue that social media can be used to broadcast "hoax

messages", which can "hide among the stream of natural messages and be accepted…".

E. Methodologies for Analyzing Blog Data

A review of the literature shows a variety of methodologies for conducting data analysis on blog data. Al-Ani et al. [22] utilized the technique of topic modeling on blog data to ascertain how blog topics changed over time between 2004 and 2011. The topic modeling technique revealed time-specific themes in the blog data that could then be compared with public events and citizen actions such as protests [22]. Berendt et al. [26] used text and graph mining to analyze blog data. The text mining revealed the word themes in the data; whereas the graph mining exposed the connections between the bloggers in terms of their use of these words/themes [26]. In addition to longitudinal content analysis, Valentini et al. [14] applied sentiment analysis to blog data wherein they attempted to assign the sentiment values of neutral, positive, or negative to each blog post. Their analysis revealed interesting time-relevant topics and associated sentiments, such as a strong lack of trust in public and private governing entities surrounding political issues [14].

F. Migration Potential Resulting from Socio-Political Crisis?

Google [28] created an interactive data visualization for tracking the interest in migrating to various countries of destination from specific countries of origin. Of particular interest is the data visualization showing a ranking of 1 to 211 based on user search counts about migrating to the United States from within a particular country. For example, Venezuela is "ranked 105 out of 211 countries for Google searches for immigration to the United States from 2014-2015" [28]. According to numerous sources, Venezuelans are protesting against their government due to such reasons as unemployment or low income, lack of access to basic needs such as food and medicine for themselves and their children, and political corruption [29, 11, 30]. Although the Google search data visualizations [28] precludes many of the Venezuelan issues that have been recently reported, one can still ask the question of whether or not there is potential for Venezuelans to migrate away from the many problems that they have been experiencing in their specific socio-political crisis? An analysis of numerous active Venezuelan-specific blogs was conducted to provide insights.

III. METHODOLOGY

In this section we discuss the methodology for this study including how the data set was obtained and the data analysis methods used.

A. Data Collection, Cleaning, and Indexing

Using the Blogtrackers tool, three steps were executed in order to crawl and collect the data from an identified set of Venezuelan blogs: (1) exploring the blog site, (2) crawling the blog site, and (3) cleaning and storing the data in a database for analysis and retrieval (Fig. 1). Hussain et al. provide a detailed explanation of the mechanics of the

Blogtrackers tool in "Analyzing the Voices during European Migrant Crisis in Blogosphere" [31].

*1) Exploring the blog site: S*everal blog sites were identified that specifically discussed issues relevant to Venezuela. Subsequently, each site was explored to determine whether their structure was ideal for use with the Blogtrackers tool. It was also important that the blog continually be focused on Venezuelan topics and contain specific metadata attributes for each post such as author, title, and date.
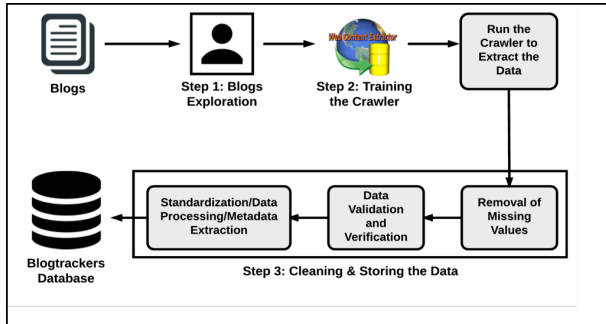

Figure 1. Data collection process.

2) Crawling the data: The Web Content Extractor (WCE) tool [32] (Fig. 2) was then used to collect data from each blog site. Once the crawler is set up, the tool begins from a set of seed URLs—the blog sites' home pages—and advances through each blog post to extract all of the desired attributes.
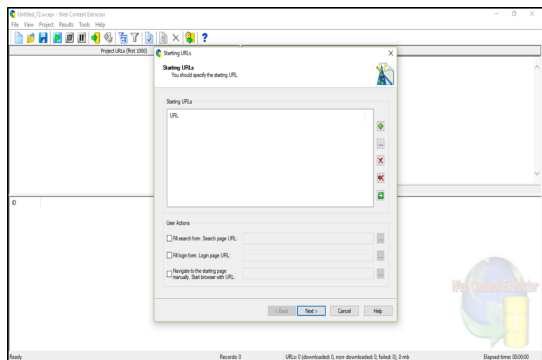

Figure 2. Web Content Extractor.

3) Cleaning and Storing the Data: A three-step cleaning process was used. (1) Clean from within WCE by deleting empty fields and advertisement URLs. (2) Clean with SQL queries to select validated and verified data. (3) Clean with a script to standardize attributes, extract metadata, sentiments, and outbound URLs.

B. Analyzing the Data with Blogtrackers

Blogtrackers is a tool designed to explore the blogosphere and gain insights about events and how these events are perceived in the blogging community [33]. After

setting up a Venezuelan blog tracker, five features of the Blogtrackers tool were used to analyze the resultant dataset.

1) Posting Frequency

The "Posting Frequency" feature was utilized to identify any unusual patterns in the blog postings. This aids in detecting real-time events that interested the blogging community. The user can click on any data point on the graph to get a detailed list of the named-entities mentioned in blog posts during that time-period. This feature also displays a list of active bloggers with number of posts. Fig. 3 shows the posting frequency for Venezuelan blogs from 2003 to 2017.
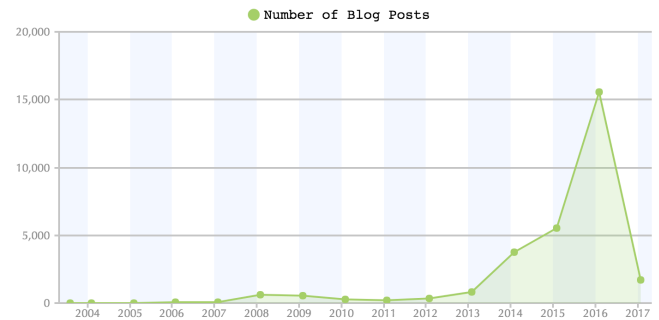

Figure 3. Change in Venezuelan blog posting trends from 2003 to 2017.
Reader is encouraged to use digital version for better readability of images, labels, legends, and captions.

2) Keyword Trends

The "Keyword Trends" feature was used to provide an overall trend for keywords of interest. It helps track changes in topics of interest in the blogging community. The user can select any data point on the trendline to view all the blogs and a network of co-occurring named-entities. Fig. 4 shows the keyword trends related to the ongoing Venezuelan socio-political crisis.
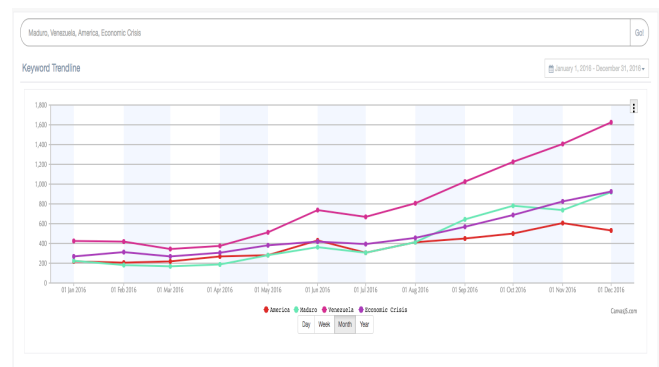

Figure 4. Trends for keywords "Venezuela", "Maduro", "America" and "Economic crisis" for 2016.
Reader is encouraged to use digital version for better readability of images, labels, legends, and captions.

3) Sentiments and Tonality

The "Sentiments and Tonality" feature was used to display the trend of positive and negative sentiments of blogs for a selected time-period (Fig. 5).
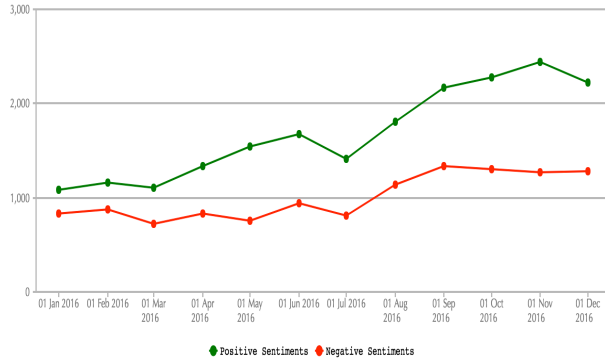
Figure 5. Venezuelan blog sentiment trends for 2016.
Reader is encouraged to use digital version for better readability of images, labels, legends, and captions.

The sentiment and tonality features are defined by Pennebaker et al. [34, 35] and as calculated by the Linguistic Inquiry and Word Count (LIWC) software [36]. Additionally, a data analyst can drill down by clicking on any point of interest and view radar charts (Fig. 6), which display tonality attributes such as personal concerns, time orientation, core drives, and cognitive process.

Figure 6. Tonality of two random Venezuelan blog posts.
Reader is encouraged to use digital version for better readability of images, labels, legends, and captions.

4) Influence

The "Influence" feature was used to identify the influence of a blogger or a post. Agarwal et al. discussed the concept of bloggers' *influence* [6]. The "Influence" feature of the Blogtrackers tool can display the influence trends over time for the top 5 influential bloggers (Fig. 7). Clicking on a point on the trend line allows a deeper dive into the data by displaying the most influential posts for that period. Additionally, a user can explore the content themes of active-influential, inactive-influential, active-non influential, and inactive-non influential bloggers.
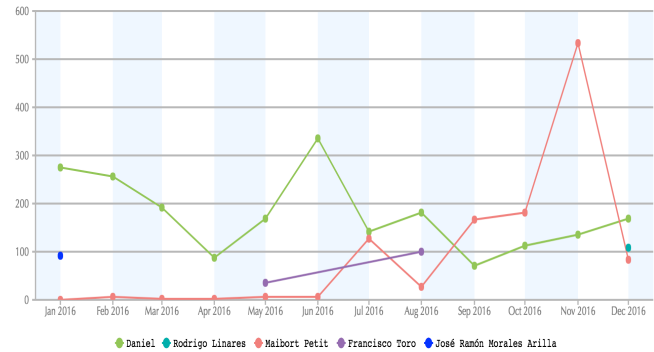

Figure 7. Influence trend for top 5 bloggers for 2016.
Reader is encouraged to use digital version for better readability of images, labels, legends, and captions.

5) Additional Blog Information

The "Additional Blog Info" feature was used to provide additional information about a specific blog. A dashboard-like screen is presented to the user, revealing the posting trends and sentiments of the selected blogs, as well as a list of the underlying URLs and domains. Fig. 8 shows this feature being used to look at a specific blog from the database called, "Caracas Chronicles". At a glance, we can see some interesting things about this particular blog; such as, it is most active on Monday's; it was most active during the month of February in 2014, December in 2015, and April in 2016; and there was more negative sentiment in 2016 than in 2015.

Figure 8. "Additional Blog Info" Blogtrackers feature for selected Venezuelan blog.
Reader is encouraged to use digital version for better readability of images, labels, legends, and captions.

IV. DISCUSSION

In this section we describe the database of Venezuelan blog posts that we collected in accordance to our above-mentioned methodology, and we discuss details from our data analysis methods.

A. Venezuelan Blog Database

To assess whether Venezuelan blogs were discussing issues related to the economic crisis in Venezuela, 40 blog sites were identified. The blogs were found using simple manual search techniques on various platforms, such as google.com, blogsearchengine.org, and fastblogfinder.com. The blogs were reviewed by our research team to ensure that they fit the structure required by the WCE and the Blogtrackers tool. A final dataset of 29,493 blog posts was obtained between August 27, 2003 and March 26, 2017. A total of 177,870 links were extracted (120,296 being distinct links) from 13,590 domains and 749,829 entities. The post sentiments were also extracted. Table I provides the language distribution for this dataset.

TABLE I. LANGUAGE STATISTICS

Language	Blogs	Blog Posts
Spanish	23	16,916
English	29	12,490
Italian	3	51
French	3	3
Portuguese	2	2
German	1	2
Breton	1	1
Catalan	1	1
Polish	1	1

B. Posting Frequency of Venezuelan Blogs

Fig. 9 shows a more detailed view of blog posting frequency from January 2015 to March 2017, which indicates a continuous increase. Specifically, blogging activity increased drastically between March 2016 and January 2017.
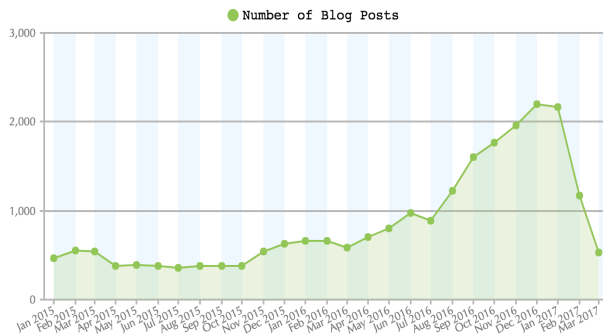

Figure 9. Change in Venezuelan blog posting frequency from January 2015 to May 2017.
Reader is encouraged to use digital version for better readability of images, labels, legends, and captions

This increase corresponds to news reports of hundreds of thousands of people beginning to take to the streets in protest during early September of 2016 [37]. Between the beginning of January, 2015 and the beginning of March, 2017, blogging frequency can be said to have notably risen by the beginning of June, 2016 (increasing by about 50%) (Fig. 9).

C. Keyword Trendlines for Venezuelan Blog

To further assess the extent of the impact of the crisis on Venezuelan citizens, the dataset was searched for quality of life keywords such as "need food", "need water", "need petrol", "need medicine", "high prices", and "inflation". The resultant keyword trendlines indicated that the occurrence of these quality of life factors fluctuated over time (Fig. 10). The occurrence of English keywords related to quality of life such as "need food", "food shortage", "high prices", "need water", "need petrol", "need medicine", and "inflation" all began to notably increase beginning around January 1, 2012 (Fig. 10).
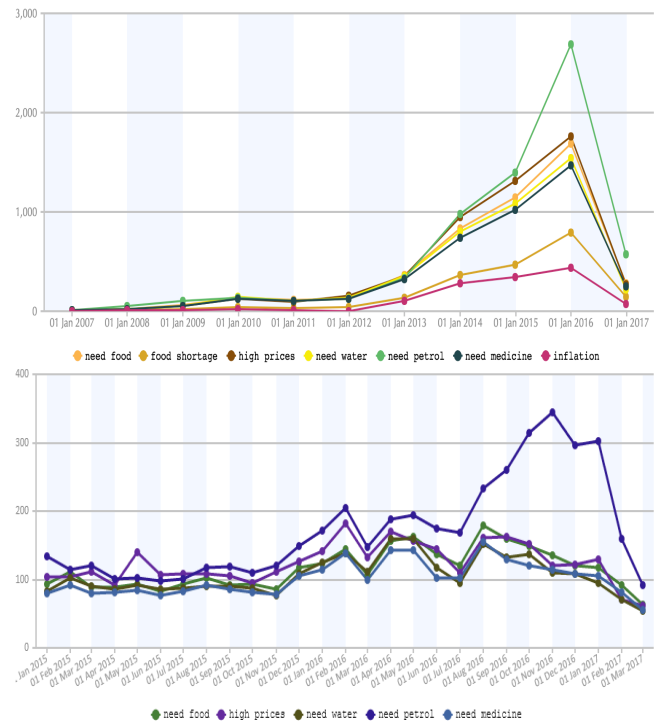

Figure 10. English keyword trendlines for various quality of life keywords over time.
Reader is encouraged to use digital version for better readability of images, labels, legends, and captions.

Trading Economics data shows that beginning January, 2012, crude oil prices began to decline and the Venezuelan monetary exchange rate plummeted [39, 40]. Of the above-mentioned keywords, "need petrol" was the most prevalent since around mid-2013 through January 1, 2017, spiking dramatically in early January of 2016. On February 18, 2016, CNN Money reported that on the previous day, Venezuelan President Nicholas Maduro announced a 6,000% price increase for gasoline/petrol in Venezuela [41]. Additionally, of these keywords, "high prices", "need food",

"need water", and "need medicine" were the next highest ranking, respectively. This ranking order of occurrence remained the same from early January 2014 through the end of 2016. Beginning early January 2016, however, the occurrence of all of these keywords began to decline through early January of 2017. A broader look from January 2015 through March 2017 reveals a spike in the occurrence of the keyword "high prices" in early May 2015 and again in early February 2016. The spike in early May 2015 may correspond to incidents of raids on Venezuelan markets by "black-market foot soldiers" beginning in early April [42]. This broader view also shows a spike in the occurrence of "need petrol" in early February 2016. The count declines from the rest of February 2016 to early March 2016, then begins to climb again through early May 2016. There was a drastic increase beginning early July 2016, which peaked in early November 2016. There was steady decline from early January 2017 to March 2017, where the occurrence of the keyword "need petrol" returned to levels closer to the other noted keywords. The keyword "need food" broke out to become the second-highest occurring keyword beginning early July 2016 (recall the NYT report of Venezuelans ransacking stores). It held this position through March 2017, closely followed by "high prices", which exceeded it slightly and briefly in early January 2017. The occurrence of "need medicine" also rose significantly from early July 2016 through early August 2016. The occurrence of "need water" rose to the fourth-highest position from early September 2016 through late October 2016.

Trendlines for the Spanish equivalents for these quality of life keywords were also generated (Fig. 11). For example, similar to the increase in blog posting frequency, the quality of life keywords trends figures each show a large increase in frequency beginning near early September of 2016 [37, 43]. The occurrence of the Spanish equivalents of many of these keywords related to quality of life such as "escasez de alimentos", "necesito agua", "inflacion", "cuidado de la salud", "necesito comida", "altos precios", "necesito medicina", and "necesito gasolina" also saw variation over time (Fig. 11). The occurrence of the Spanish keyword "inflacion" (inflation) did not register on the graphical output for January 1, 2003 to January 1, 2017, and "necesito gasolina" (need gasoline) barely registered. The occurrence of "escasez de alimentos" (food shortage) and "cuidado de la salud" (health care) mirrored each other, beginning to increase in early January 2013, increasing further sharply in early January 2015, peaking in late December 2016, and dropping sharply through early January 2017. "Necesito agua" (need water) consistently held the second-highest occurrence count notably from early January 2014 through early January 2017, with a breakout spike between early January to late December 2015. Also through this early January to late December 2015 timeframe, "altos precios" (high prices), "necesito medicina" (need medicine), and "necesito comida" (need food) held the next highest count rankings, respectively. "Necesito agua" (need water) saw a

spike in occurrence count between early January 2008 and early January 2009. A broader look at the occurrence of the Spanish keywords from January 2015 through March 2017 reveals that "cuidado de la salud" (health care) and "escasez de alimentos" (food shortage) consistently mirrored each other and held the top rank for most occurring of the quality of life keywords. "cuidado de la salud" and "escasez de alimentos" occurred at least 50% more of the time than the others from early January to early March 2015, then beginning again from mid-November 2015. Although experiencing a short decline from early June to early July 2016, the count of these two keywords continued to rise through their peak in early December 2016. They began to decline sharply from early January 2017, returning to near January 2015 levels by early March 2017. "Necesito agua" began in a breakout rise in early May 2016, with a mild spike in early June 2016 (corresponding to the NYT's story of ransacking stores), another spike in early December 2016, then a slow decline through early March 2017. "Altos precios", "necesito medicina", "inflacion", and "necesito comida" (in that apparent ranking order) only began to register on the data visualization from around early October 2016 through around early January 2017. "Necesito gasolina" barely registered on the data visualization in this 14-month period except from early November 2016 to around mid-February 2017.
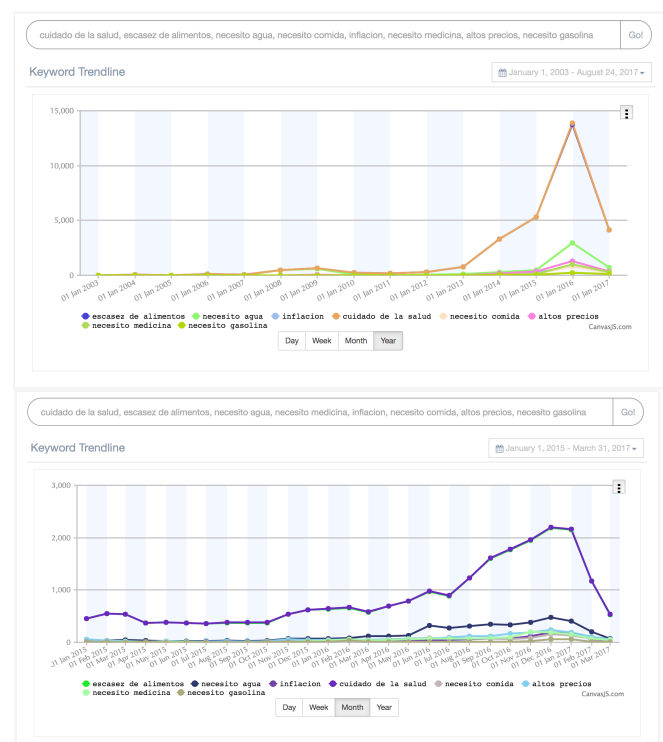

Figure 11. Spanish keyword trendlines for various quality of life keywords over time.
Reader is encouraged to use digital version for better readability of images, labels, legends, and captions.

These quality of life keywords may represent a set of motivational factors leading to an interest in Venezuelan citizens migrating away from Venezuela. To explore this suggestion, we generated several trendlines for keywords such as "immigrate", "migrate", "emigrate", and "leave Venezuela" (Fig. 12).
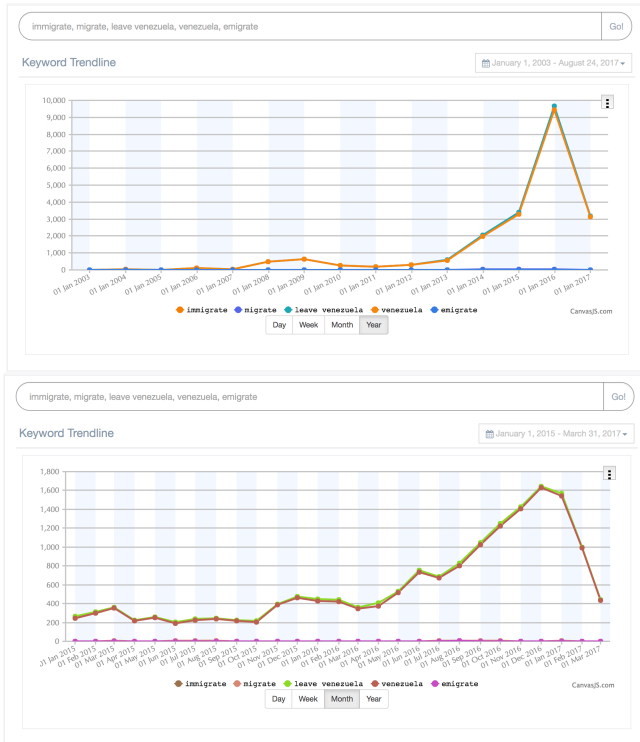

Figure 12. English keyword trendlines for various keywords that possibly indicate migration interest over time.
Reader is encouraged to use digital version for better readability of images, labels, legends, and captions.

The frequency of occurrence of the keyword, "leave Venezuela" seems to drastically increase between the end of 2015 and beginning of 2016. The keyword "leave Venezuela" began to increase dramatically in occurrence beginning early January 2013 through its peak of almost 10,000 posts in late December 2015, with the most notable rise being between early January 2015 and late December 2015. According to a report by United Press International, in 2016, 75% of the Venezuelan population lost an average of 19 pounds due to the effects of the economic crisis that was satirically referred to as the "Maduro Diet" [44]. By early January 2017, occurrences of "leave Venezuela" dropped down to about 3,000 (Fig. 12). The bottom chart of Fig 12 kind of zooms in to show that the peak in occurrence of the keyword "leave Venezuela" actually occurred in early December 2016.

Trendlines for the Spanish equivalents for these migration keywords were also generated (Fig. 13). It was revealed that the keywords "leave Venezuela" and "Venezuela" were almost identical, indicating that every time Venezuela was mentioned it was about leaving or migrating from Venezuela. As with the quality of life indicators, these migration-related keywords trends figures show a large increase in their frequency beginning in 2016, especially near early-September. Of the Spanish equivalents of English keywords that may be indicative of a desire to flee Venezuela, "salir de Venezuela" (leave Venezuela) was the only one that notably registered on the data visualization between early January 2003 and early January 2017 (Fig. 13).
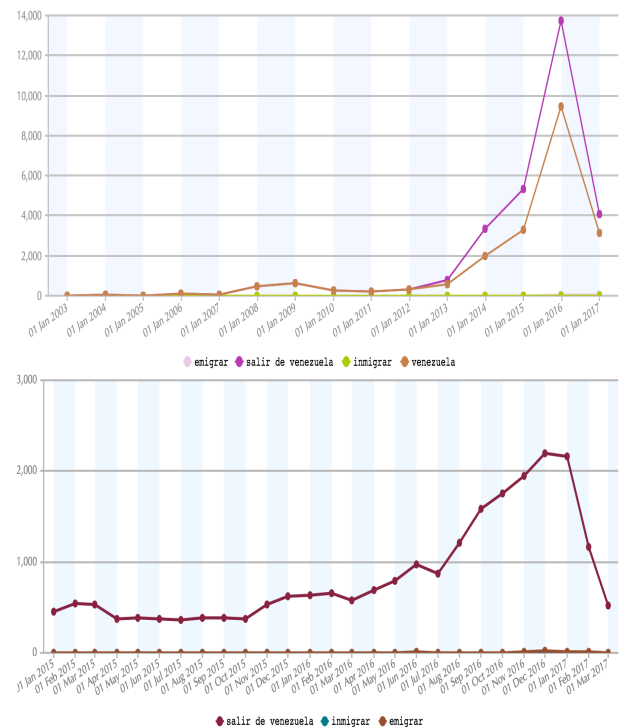

Figure 13. Spanish keyword trendlines for various keywords that possibly indicate migration interest over time.
Reader is encouraged to use digital version for better readability of images, labels, legends, and captions.

The others, "emigrar" (emigrate) and "inmigrar" (immigrate), barely registered. Occurrences of the keyword "salir de Venezuela" began to increase in early January 2012 (recall the Trading Economics data showing the decline in both crude oil prices and the Venezuelan monetary exchange rate), peaking in late December 2016 near 14,000, then beginning to decrease sharply through late December 2016 or early January 2017. According to a report by CNN Money, there was a 160% increase from 2015 to 2016 in Venezuelans seeking U.S. asylum [45]. The most notable increases in the occurrence of the keyword "salir de Venezuela" (each at different rates) were between early January and late December 2013, early January 2014 and early January 2015, and early January to late December 2015. Between early January 2015 and early March 2017, "salir de Venezuela" occurred at least 50% more of the time than did the keywords "inmigrar" (immigrate) and "emigrar" (emigrate). At its peak for this date range, "salir

de Venezuela" occurred over 200% more often than "inmigrar" and "emigrar". There was a slight decline in the occurrence of the keyword "salir de Venezuela" from early June to early July 2016. Subsequently, however, there was a resumption of the sharp rise of the occurrence of "salir de Venezuela" in early July 2016 (again, recall the NYT's story of the ransacking of Venezuelan stores). The keyword "emigrar" (emigrate) began to register slightly more on the data visualization from around early November through around mid-December 2016.

D. Timeline of Sentiments Among the Venezuelan Blogs

Recall that Fig. 5 shows the prevailing sentiment of the blogs for a specific time period. Fig. 14 provides a more detailed view of how the sentiment of the Venezuelan blogs has changed over time. The graphical output appears to show a sharp shift to a prevailing positive sentiment in early January 2016. This could correspond to the Venezuelan leader, Nicholas Maduro, declaring a 60-day economic emergency for the region in order to give himself power "to pay for welfare services and food imports" [43]. Subsequently, however—and although many blog posts were identified as having a negative sentiment—the prevailing sentiment of the blogs has been positive for the past two years. This prevailing positive sentiment seems counterintuitive, as we expected to see a prevailing negative sentiment due to the ongoing Venezuelan economic crisis and continued reports of Venezuelan citizens protesting in the streets. Therefore, we believe that the concept of sentiment with regard to this dataset needs to be addressed further in future work involving a more detailed dataset and a possible revision of how sentiment is calculated. Consequently, for this paper, we can only conclude that the graphical output at this time does not seem to reveal any significant differences over time with regard to the concept of sentiment, and appears to only fluctuate along with the recorded count of blog posts. Regardless, the visualization shows that between early January 2015 and early March 2017, the overall sentiment of the Venezuelan blog posts was predominantly positive (there were more blog posts rated "positive" than there were those rated "negative") (Fig. 14). The difference (spread) between the number of Venezuelan blog posts rated "positive" and the number of blog posts rated "negative" was relatively small (from about 50 to 150) from early January through around mid-November 2015. Although the number of blog posts rated "positive" and the number of blog posts rated "negative" both continued to increase between mid-November 2015 and early January 2017, the spread widened most notably between early April and early June 2016, then again between early September 2016 through early January 2017. Between mid-year 2007 and mid-year 2009, there were more Venezuelan blog posts rated "positive" than there were blog posts rated "negative". Between early January 2010 and early January 2012, the number of blog posts rated "positive" and the number of blog posts rated "negative"

were roughly equal. In early January 2013, the spread (difference) between the number of blog posts rated "positive" and the number of blog posts rated "negative" began to widen, with more being "positive". Again, this seems odd in light of the fact that Venezuela entered into a deep recession beginning January 2014 [46].

Although fluctuating in size, this positive-dominated spread continued through at least early January 2017. The bottom chart of Fig. 14 shows the number of blog posts rated as having a "positive" sentiment relative to the number of blog posts rated as having a "negative" sentiment for each day from May 2, 2016 to June 30, 2016.
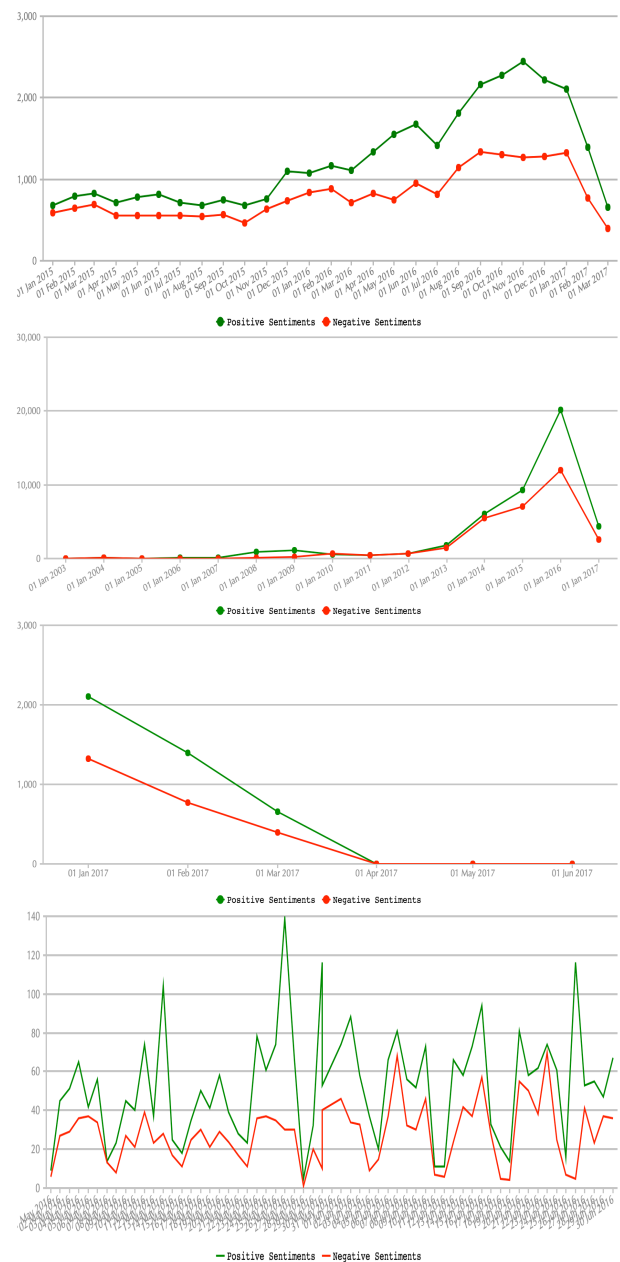

Figure 14. Change in sentiment of Venezuelan blogs over time.
Reader is encouraged to use digital version for better readability of images, labels, legends, and captions.

During this time range, the number of blog posts rated "positive" always exceeded the number of blog posts rated "negative", although only slightly for at least seven dates. On at least four dates, the number of blog posts rated "positive" sharply exceeded the number of blog posts rated "negative"; for example, on May 13, May 26, May 30, and June 26. Reuters reported that on May 13, 2016, Maduro declared a 60-day "state of emergency" [47]. News reports for May 26, 2016 also include topics related to the "state of emergency" declaration; Reuters reported that Venezuelan opposition to Maduro, essentially, warned him against corruption related to contracts with foreign companies [48]. On May 30, 2016, BBC News reported that eleven people were killed in a recession-related shooting in Trujillo state [49]. Finally, on June 26, 2016, NPR broadcasted a story on its "All Things Considered" radio program discussing how Venezuelans are running out of food and medicine [50]. At least one date stands out on the graph where the number of blog posts rated "negative" rose to almost reach the number of blog posts rated "positive"—June 23, 2016. The Washington Post reported that on that date (June 23, 2016) the Organization of American States met and argued for a recall referendum against Maduro [51]. Curiously, the bottom-most chart of Figure 14 seems to show that on May 28, 2016 there were barely any blog posts. This may be due to the system of scheduled electricity blackouts that the Venezuelan government was implementing as reported by The NYT on this date [52].

E. Tonality of Venezuelan Blogs

Recall that Fig. 6 displays the feature of Blogtrackers that shows tonality attributes of individual blog posts. For example, for two random Venezuelan blog posts in our dataset, the predominant personal concerns were "Work" and "Money". The predominant time orientation was "Present focus". The predominant attribute for core drives and needs was "Power" (and to a lesser extent "Achievement"). The predominant cognitive process was not as clear, varying among that of "Differentiation", "Tentativeness", and "Cause". The predominant summary variable was "Analytical Thinking". Finally, the predominant sentiment/emotion for this timeframe was "Anger". We did not analyze the tonality feature further for this paper, but believe that the concept should be examined in future work.

F. Influential Venezuelan Bloggers

Recall that Fig. 7 shows the top 5 influential bloggers for a specified time period. Blogtrackers calculates influence based on the definitions and algorithms developed by Agarwal et al. [25]. Fig. 15 is another example of the feature, using January 2015 to March 2017 as the selected time-period of analysis. One blogger, "Daniel", was consistently more influential than other bloggers [53]. We did not analyze this feature further for this paper, but believe that the concept of influence within the blogosphere should

be examined in future work. With fluctuating levels of influence, between late December 2015 and early March 2017, the five most influential authors of the Venezuelan blogs in our database were "Francisco Toro", "mesaredondacontracomunistabg", "Juan Cristobal Nagel", "Manuel Madrid" and "Daniel" (Fig. 15). "Manuel Madrid" only seemed to be active during early December 2015. "Francisco Toro" apparently didn't begin activity until early June 2015, and ceased being active in early January 2017. "Juan Cristobal Nagel" apparently ceased being active in early June 2016 (again, this date corresponds to the NYT's story of Venezuelans ransacking stores). The most consistent and predominantly influential blogger was "Daniel". "Daniel" was notably the most-influential blogger during at least three key date categories: early March 2015, early December 2015, and early June 2016.
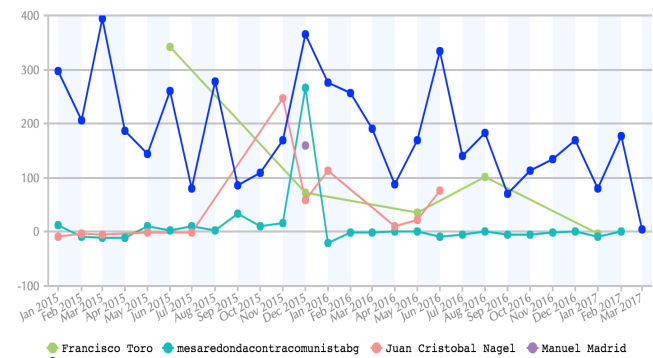

Figure 15. Influence trend of top 5 Venezuelan bloggers from January 2015 to March 2017.
Reader is encouraged to use digital version for better readability of images, labels, legends, and captions.

G. Socio-politico-economic Awareness

One event can be said to have corresponded to numerous spikes (or otherwise noticeable changes in various data categories—early June of 2016. June 19, 2016 is when news reports emerged of Venezuelans ransacking stores out of desperation stemming from food shortages and high prices. This event seemed to spark bloggers to increase their posting frequency. This event also seemed to cause bloggers to increase their use of the keywords "necesito agua" (need water). Additionally, and although the data show a seeming decline from early June to early July of 2016, the use of the keyword "salir de Venezuela" (leave Venezuela) quickly resumed in frequency in early July of 2016; perhaps communicating an increasing level of frustration with the economic conditions. This pattern with the frequency of the keywords "salir de Venezuela" comes on the heals, however, of another data pattern that popped out in the visualizations that show that the frequency of the use of the keywords "cuidado de salud" (healthcare) and "escasez de alimentos" (food shortage) experienced a short decline from early June to early July of 2016. Perhaps this means that instead of blogging about these two topics, the bloggers were discussing the possibility of migration. The data also showed

that the number of blog posts rated as having a positive sentiment began to decline around the time of this event. In fact, the number of blog posts rated as having a negative sentiment rose to almost reach the number of blog posts rated as having a positive sentiment around June 23, 2016. Another characteristic of the data that seemed to correspond to this event is that one of the top five most influential bloggers, "Juan Cristobal Nagel" apparently ceased being active in early June 2016, which, although speculation, causes one to wonder if these harsh economic conditions were behind his exit from the Venezuelan blogosphere. Ultimately, however, these patterns in the data seem to suggest that indeed bloggers, especially the Spanish-language bloggers, are raising awareness of these socio-political events.

V.   CONCLUSION & FUTURE WORK

This paper extends the work by Mead et al. that established the basis for using blog analysis for studying socio-political awareness. This approach remains novel in that few researchers have specifically focused on analyzing blogs, and instead focus on other social media platforms, such as Twitter, Facebook and YouTube. As a detailed example, this research showed that Venezuelan blogs are being used to disseminate socio-political information in an attempt to increase awareness of events and sentiments during the Venezuelan economic crisis. Our analysis showed that the frequency and content of blog posts changed over time, reflecting changes in the socio-politico-economic landscape of the region—such as protests and other news events, the documented decline in quality of life factors such as the need for food and medicine, and expressed interest in migration away from Venezuela. The sentiment of the blogs seemed to change over time as well, but the overall sentiment analysis was inconclusive and seemed counterintuitive and therefore the concept needs to be addressed further in future work. For example, the number of blog posts that were rated as having a positive sentiment far exceeded the number of blog posts rated as having a negative sentiment even when the dates revolved around significant events such as on June 26, 2016, which was only a few days after news reports emerged of Venezuelans ransacking stores out of desperation. We believe, however, that blog analysis—with Blogtrackers and other tools—can continue to be used to gauge socio-political awareness of important issues among various populations. This paper continues to develop the stage for future work using Blogtrackers and other natural language processing tools and techniques for blog analysis as a possible approach for anticipating events (e.g., protests, migration, refugee scenarios). Future work may also include further analysis of the concepts of blog tonality and blogger influence. Specifically, further understanding blogger influence is of great importance due to the emergent concepts of fake news, opinion manipulation, and disinformation campaigns. Broadly speaking, however, this particular study sheds a spotlight on the blogosphere's role in assessing situation awareness of a region engulfed in socio-politico-economic crisis. We believe that the information derived from monitoring blogs can provide actionable insights to local emergency responders (i.e., real-time blog monitoring), and humanitarian assistance organizations and policy decision makers (i.e., based on short- and long-term trend analysis).

REFERENCES

[1] E. L. Mead, M. N. Hussain, M. Nooman, S. Al-khateeb, and N. Agarwal, "Assessing Situation Awareness through Blogosphere: A Case Study on Venezuelan Socio-Political Crisis and the Migrant Influx," The Seventh International Conference on Social Media Technologies, Communication, and Informatics (SOTICS 2017), 2017, Oct. 8, 2017, Athens, Greece. ISSN: 2326-9294

[2] D. Gillmor, "We the Media: The Rise of Citizen Journalists," National Civic Review, vol. 93 (3), pp. 58–63, 2004, doi:10.1002/ncr.62.

[3] Z. Tufekci and C. Wilson, "Social media and the decision to participate in political protest: Observations from Tahir Square," Journal of Communication, vol. 62 (2), pp. 363-379, 2012, doi: 10.1111/j.1460-2466.2012.01629.x.

[4] H. S. Du and C. Wagner, "Learning with weblogs: An empirical investigation," presented at the 2005 Proceedings of the 38th Annual Hawaii international Conference on System Sciences, pp. 7b-7b, Jan. 24, 2005, doi: 10.1109/HICSS.2005.387.

[5] Y. Xie and P. Sharma, "Students' lived experience of using weblogs in a class: An exploratory study," Association for Educational Communications and Technology (AECT 2004), pp. 839-846, 2005, Accession Number: ED485009.

[6] N. Agarwal, H. Liu, L. Tang, and P. S. Yu, "Identifying the influential bloggers in a community," Proc. 2008 International Conference on Web Search and Data Mining (WSDM 2008), pp. 207–218, ACM, 2008, ISBN: 9781595939272.

[7] Blogtrackers, "Analyze anything about blogs," University of Arkansas at Little Rock, 2017. [Online]. Available from: http://blogtrackers.host.ualr.edu/ 2018.05.26

[8] D. T. Cain, "Twituational awareness: gaining situational awareness via crowdsourced# disaster epidemiology." Diss. Monterey California: Naval Postgraduate School, Sep. 2013. [Online]. Available from: https://calhoun.nps.edu/bitstream/handle/10945/37594/13Sep_Cain_Daniel.pdf?sequence=1 2018.05.26

[9] A. R. Curtis, "From Arab Spring to Shahbag: The role of social media in terms of national crisis," Journal of Mass Communication Journalism, vol. 5 (2), pp. 1-3, 2015, doi: 10.4172/2165-7912.1000241.

[10] M. Golden Pryor, M. Wulf, W. Alanazi, N. Alhamad, and O. Shomefun, "The Role of Social Media in Transforming Governments and Nations," International Journal of Business & Public Administration, vol. 11 (1), pp. 19-30, Sum. 2014, ISSN: 15474844.

[11] V. Karagiannopoulos, "The role of the internet in political struggles: Some conclusions from Iran and Egypt," New Political Science, vol. 34 (2), pp. 151-171, May 2012, doi: 10.1080/07393148.2012.676394.

[12] J. B. Lim, "Engendering civil resistance: Social media and mob tactics in Malaysia," in International Journal of Cultural Studies, vol. 20 (2), Mar. 2017, doi: 10.1177/1367877916683828.

[13] N. Pang and D. Goh, "Can blogs function as rhetorical publics in Asian democracies? An analysis using the case of Singapore," Telematics and Informatics, vol. 33 (2), pp. 504-513, 2016, doi: 10.1016/j.tele.2015.08.001.

[14] C. Valentini and S. Romenti, "Blogging about crises: The role of online conversations in framing Alitalia's performance during its crisis," Journal of Communication Management, vol. 15 (4), pp. 298-313, 2011, doi: 10.1108/13632541111183398.

[15] L. J. Van Leuven, "Optimizing citizen engagement during emergencies through use of Web 2.0 technologies," Thes. Monterey California Naval Postgraduate School, Mar. 2009. [Online]. Available from: https://calhoun.nps.edu/bitstream/handle/10945/4819/09Mar_Van_Leuven.pdf?sequence=1 2018.05.26

[16] F. H. Abdullah, "The Role of Social Network Platform in Egyptian's Political Upheaval in January 2011," International Journal of Social Sciences & Educational Studies, vol. 3 (2), pp. 94-108, Dec. 2016, ISSN: 2409-1294.

[17] B. Birregah, T. Top, C. Perez, E. Chatelet, N. Matta, M. Lemercier, and H. Snoussi, "Multi-layer crisis mapping: a social media-based approach," Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE 2012), IEEE, Jun. 2012, pp. 379-384, doi: 10.1109/WETICE.2012.47.

[18] Y. Liu, P. Piyawongwisal, S. Handa, L. Yu, Y. Xu, and A. Samuel, "Going beyond citizen data collection with mapster: a mobile+ cloud real-time citizen science experiment," e-Science Workshops, IEEE Seventh International Conference, (eScienceW 2011), IEEE, Dec. 2011, pp. 1-6, doi: 10.1109/eScienceW.2011.23.

[19] A. M. MacEachren, A. Jaiswal, A. C. Robinson, S. Pezanowski, A. Savelyev, P. Mitra, and X. Zhang, "Senseplace2: Geotwitter analytics support for situational awareness," IEEE Conference on Visual Analytics Science and Technology (VAST 2011), pp. 181-190. IEEE, 2011, doi: 10.1109/VAST.2011.6102456.

[20] J. Yin, A. Lampert, A. Cameron, R. Robinson, and P. Power, "Using social media to enhance emergency situation awareness," IEEE Intelligent Systems, vol. 27 (6), pp. 52-59, Feb. 2012, doi: 10.1109/MIS.2012.6.

[21] J. Zhu, F. Xiong, D. Piao, L. Liu, and Y. Zhang, "Statistically modeling the effectiveness of disaster information in social media," Global Humanitarian Technology Conference (GHTC 2011), pp. 431-436. IEEE, Oct. 2011, doi: 10.1109/GNTC.2011.48.

[22] B. Al-Ani, G. Mark, J. Chung, and J. Jones, "The Egyptian blogosphere: a counter-narrative of the revolution," Proc. ACM 2012 conference on Computer Supported Cooperative Work, pp. 17-26. ACM, 2012.

[23] S. I. Yuce, N. Agarwal, and R. T. Wigand, "Women's Right to Drive: Spillover of Brokers, Mobilization, and Cyberactivism," International Conference on Social Computing, Behavioral-Cultural Modeling, and Prediction

(SBP 2015), vol. 9021, pp. 232-242, Mar. 2015, doi: 10.1007/978-3-319-16268-3_24.

[24] P. P. Y. Leong, "Political Communication in Malaysia: A Study on the Use of New Media in Politics," JeDEM-eJournal of eDemocracy and Open Government, vol. 7 (1), pp. 46-71, 2015, ISSN: 2075-9517.

[25] R. Goolsby, L. Shanley, and A. Lovell, "On cybersecurity, crowdsourcing, and social cyber-attack," Office of Naval Research, Wilson Center, Commons Lab, policy memo series, vol. 1, 2013. [Online]. Available from: http://www.dtic.mil/get-tr-doc/pdf?AD=ADA580185 2018.05.26

[26] B. Berendt and R. Navigli, "Finding Your Way through Blogspace: Using Semantics for Cross-Domain Blog Analysis," AAAI Spring Symposium: Computational Approaches to Analyzing Weblogs, 2006, pp. 1-8.

[27] N. Fearn, "When bots go roque," International Data Group (IDG) Connect. Feb. 17, 2017. [Online]. Available from: http://www.idgconnect.com/blog-abstract/24495/when-bots-rogue 2018.05.26

[28] Google, "Google Trends: Migrant Crisis-A human tragedy," 2015. [Online]. Available from: https://trends.google.com/trends/story/US_cu_iQ-fCFEBAABHhM_en 2018.05.26

[29] Fox News, "Study: Venezuelans lost 19 lbs. on average over past year due to lack of food," Feb. 20, 2017. [Online]. Available from: http://www.foxnews.com/world/2017/02/20/study-venezuelans-lost-19-lb-on-average-over-past-year-due-to-lack-food.html 2018.05.26

[30] L. Melendez, "Encovi 2016: 74% of Venezuelans lost more than 8 kilos of weight last year," Feb. 18, 2017. [Online]. Available from: http://runrun.es/rr-es-plus/297797/encovi-2016-74-de-los-venezolanos-perdio-mas-de-8-kilos-de-peso-el-ano-pasado.html 2018.05.26

[31] M. N. Hussain, K. Bandeli, M. Nooman, S. Al-khateeb, and N. Agarwal, "Analyzing the Voices during European Migrant Crisis in Blogosphere," 2nd International Workshop on Event Analytics using Social Media Data associated with The 11th International AAAI Conference on Web and Social Media (ICWSM 2017), May 15-18, 2017, Montreal, Canada.

[32] Newprosoft, "Web Content Extractor," 2017. [Online]. Available from: http://www.newprosoft.com/web-content-extractor.htm 2018.05.26

[33] N. Agarwal, H. Liu, L. Tang, and S. Y. Philip, "Modeling blogger influence in a community," Social Network Analysis & Mining, vol. 2 (2), pp. 139–162, Jun. 2012, ISSN: 18695450.

[34] J. W. Pennebaker, R. J. Booth, and M. E. Francis, "Operator's Manual Linguistic Inquiry and Word Count: LIWC 2007," LIWC.net, Austin, Texas. [Online]. Available from: http://www.depts.ttu.edu/psy/lusi/files/LIWCmanual.pdf 2018.05.26

[35] J. W. Pennebaker, R. L. Boyd, K. Jordan, and K. Blackburn, "The development and psychometric properties of LIWC2015," Sep 15 2015. [Online]. Available from: http://liwc.wpengine.com/wp-content/uploads/2015/11/LIWC2015_LanguageManual.pdf 2018.05.26

[36] "LIWC - Linguistic Inquiry and Word Count." [Online]. Available from: http://liwc.wpengine.com/ 2018.05.26

[37] N. Casey, "Thousands March in Venezuela to Demand President's Ouster," New York Times, Sep. 1, 2016. [Online]. Available from: https://www.nytimes.com/2016/09/02/world/americas/caracas-venezuela-nicolas-maduro-protests.html?_r=0 2018.05.26

[38] N. Casey, "Venezuelans Ransack Stores as Hunger Grips Nation," New York Times, June 19, 2016. [Online]. Available from: https://www.nytimes.com/2016/06/20/world/americas/venezu elans-ransack-stores-as-hunger-stalks-crumbling-nation.html 2018.05.26

[39] Trading Economics, "Crude Oil 1946-2018," 2018. [Online]. Available: https://tradingeconomics.com/commodity/crude-oil 2018.05.26

[40] Trading Economics, "Venezuelan Bolivar 1988-2018," 2018. [Online]. Available: https://tradingeconomics.com/venezuela/currency 2018.05.26

[41] J. Wattles and P. Gillespie, "Venezuelans face a 6000% hike in gasoline price," CNN Money, February 18, 2016. [Online]. Available: http://money.cnn.com/2016/02/17/news/world/venezuela-gas-price-hike/index.html 2018.05.26

[42] M. Mogollon, "Packs of black-market foot soldiers raid Venezuela markets," LA Times, Apr. 1, 2015. [Online]. Available from: http://www.latimes.com/world/mexico-americas/la-fg-venezuela-shopping-20150401-story.html 2018.05.26

[43] BBC News, "Venezuela economy: Nicolas Maduro declares emergency," Jan. 15, 2016. [Online]. Available from: http://www.bbc.com/news/world-latin-america-35329617 2018.05.26

[44] A. V. Pestano, "Venezuela: 75% of population lost 19 pounds amid crisis," United Press International, February 19, 2017. [Online]. Available: https://www.upi.com/Venezuela-75-of-population-lost-19-pounds-amid-crisis/2441487523377/ 2018.05.26

[45] P. Gillespie, "Thousands of Venezuelans fleeing to the US," CNN Money, May 23, 2017. [Online]. Available: http://money.cnn.com/2017/05/23/news/economy/venezuela-us-asylum-refugees/index.html 2018.05.26

[46] Z. Aleem, "How Venezuela went from a rich democracy to a dictatorship on the brink of collapse," Vox, September 19, 2017. [Online]. Available: https://www.vox.com/world/2017/9/19/16189742/venezuela-maduro-dictator-chavez-collapse 2018.05.26

[47] A. Ulmer and C. Pons, "Venezuela opposition slams 'desperate' Maduro state of emergency," Reuters, May 13, 2016. [Online]. Available: https://www.reuters.com/article/us-venezuela-politics/venezuela-opposition-slams-desperate-maduro-state-of-emergency-idUSKCN0Y501X 2018.05.26

[48] Reuters Staff, "Venezuelan opposition warns foreign contracts need assembly's OK," Reuters, May 26, 2016. [Online]. Available: https://www.reuters.com/article/us-venezuela-politics/venezuelan-opposition-warns-foreign-contracts-need-assemblys-ok-idUSKCN0YH2L9 2018.05.26

[49] BBC News: Latin America, "Venezuela: Eleven shot dead in Trujillo state," BBC News, May 30, 2016. [Online]. Available: http://www.bbc.com/news/world-latin-america-36410563 2018.05.26

[50] NPR: All Things Considered, "Running Out of Food, Medicine and Patience in Venezuela," June 26, 2016. [Online]. Available: https://www.npr.org/2016/06/26/483624346/running-out-of-food-medicine-and-patience-in-venezuela 2018.05.26

[51] C. Morello, "OAS head calls for recall of Maduro to restore democracy in spriling Venezuela," The Washington Post, June 23, 2016. [Online]. Available: https://www.washingtonpost.com/world/national-security/oas-head-calls-for-recall-of-maduro-to-restore-democracy-in-spiraling-venezuela/2016/06/23/1568edcc-3974-11e6-a254-2b336e293a3c_story.html 2018.05.26

[52] N. Casey and P. Torres, "Venezuela Drifts Into New Territory: Hunger, Blackouts and Government Shutdown," New York Times, May 28, 2016. [Online]. Available: https://www.nytimes.com/2016/05/28/world/americas/venezu ela-economic-government-collapse.html 2018.05.26

[53] D. Duquenal, "Venezuela News And Views: A blog about life under, and resisting, a dictatorship". [Online]. Available from: http://daniel-venezuela.blogspot.com 2018.05.26