

Analyzing Shift in Narratives Regarding Migrants in Europe via Blogosphere

Muhammad Nihal Hussain
Information Science
mnhussain@ualr.edu

Kiran Kumar Bandeli
Information Science
kxbandeli@ualr.edu

Samer Al-khateeb
Information Science
sxalkhateeb@ualr.edu

Nitin Agarwal
Jerry L. Maulden-Entergy Chair Professor of Information Science
nxagarwal@ualr.edu

University of Arkansas at Little Rock, Little Rock, United States

Abstract

Social media is widely used by individuals to express their views or opinions with others. Social media users leverage this platform to further their views by framing narratives and participating in online discourse. Nowadays almost all events, issues, and crises are discussed on social media. Blogs are not regulated by any authority and have no limit on the number of characters - unlike other social media platforms - which provide bloggers with a richer space of content. Blogs also serve as a platform for agenda-setting and content framing abetting development of weaponized narratives. Blogs are a good source of data for sociologists and political scientists to gain situational awareness about various events which can be achieved by tracking different opinions and political views as being shaped. In this research, we analyze blogs to study shift in narratives in blogosphere towards refugees or migrants during the migrant crisis in Europe. We use the Blogtrackers tool to analyze over 9,000 blog posts published from December 2005 to mid-March 2016. We use named-entity extraction to identify different topics and themes, then use targeted sentiment analysis to study the shift in narratives toward migrants in the blogosphere.

1 Introduction

Social media which was once considered merely as a hub for high school students to connect and share various pictures, videos, or tweets has now become a powerful medium for individuals to express their views and perspectives about various events. Social media today has risen to new heights, and it has become a primary source of news for many adults [Gri17, SG17]. While social media is mainly used for benevolent purposes, a few use this

Copyright © 2018 for the individual papers by the paper's authors. Copying permitted for private and academic purposes. This volume is published and copyrighted by its editors.

In: A. Jorge, R. Campos, A. Jatowt, S. Nunes (eds.): Proceedings of the Text2StoryIR'18 Workshop, Grenoble, France, 26-March-2018, published at <http://ceur-ws.org>

platform for deviant¹ purposes such as cyber-bullying, cyber-warfare, propaganda and fake news dissemination to influence the masses [noa18, Alb16] or disorient them to provoke hysteria among citizens. It is important to study the various narratives on social media and to gain situational awareness which should help in stemming the effect of deviant usage of social media. As the information deluges on social media, it is a challenge to understand what narratives propagate in social media and particularly on blogs.

In this research, we apply concepts of *thematic narrative analysis* [Rie93], i.e., focus on “what” is said more than “how” it is said, to identify leading narratives and how a shift in the narratives takes place in the blogosphere. We achieved this by using *sentiment trend analysis* and *keyword trend analysis*. Through a longitudinal analysis and corroborating with real world events, we further try to understand why the shift in narratives occurs. We conduct the analysis using the Blogtrackers tool [Blo18]. For this study, we considered the events during 2014-2016 migrant influx in European Union. However, the research methodology can be applied to any other real-world event [MHN⁺]. This is important given the abundant information on social media especially blogs. It is hard to know when the shift in narrative happens and what the audiences are expressing. Our analysis using the Blogtrackers tool can help social scientists identify narratives. This enables the analysts to learn what resonates with the community and if those interests and views are changing with time under the influence of exogenous factors or events.

The rest of the article is organized as follows. Section 2 provide a brief summary of the literature conducted on narrative analysis. We discuss the research methodology in section 3, and include our finding in section 4. We conclude with intended future work in section 5.

2 Literature Review

Many recent studies on narrative analysis focus on different types of events. A study conducted by Chou et al. [CHFA11] focuses on using linguistic-based narrative analysis of YouTube accounts. The analysis elucidates the common attributes of the narratives. It also identified and analyzed a list of shared thematic and linguistic characteristics.

Another study by Corman et al. [CRF12] presents how narratives are used as key features by extremists. They built a network of stories and relate these stories based on semantics. This network can be analyzed directly - the more common words, semantically, among stories the more these stories are related - or used as input to clustering algorithms to identify similar stories.

Another study by Miranda [Mir] focuses more on narrative and social media by discussing how narrative describes the past, justifies the present, and presents a vision of the future. Further, it explains how multiple interconnected narratives provide intent and justification of a strategy to different target audiences.

Another study by Ruston et al. [RCS⁺16] presents the case where narrative can be used by individuals to make sense of the world around them. It also helps individuals to quickly decide whether they believe or discard information. Further, narrative helps in shaping and expressing individual’s ideology also in executing the political functions of ideology.

Our work is different than the aforementioned works as we study blogs first and then identify narratives computationally using *sentiment trend analysis* and *keyword trend analysis*. Later, we detect the shift in narratives using the Blogtrackers tool.

3 Research Methodology

Subject matter experts used local news reporting about various organizations, entities, and events to identify 22 blog sites related to migrant issues in Europe. We used the 3-step crawling methodology described in [HOB⁺17] to setup crawlers that collect data from blogs. Our crawlers were used to extract the following data attributes: post title, author/blogger name, date of posting, number of comments, and permalinks for each post. From the seed of 22 blog sites, we collected 9,183 blog posts. The blog posts were published during the period December 2005 to mid-March 2016.

After the data was crawled, we cleaned it to remove noise (e.g., records with empty or incorrect attributes, JavaScript codes of social media plugins or advertisements, etc.), and standardized the attributes to be in a consistent format (e.g., dates in different formats like “martes, 2 de agosto de 2016”, “27 Apr 2017”). We also enriched the data by extracting sentiments using one of the “gold standard computerized text analysis”, i.e., Linguistic

¹Throughout this paper, when we mention the word “deviant” we mean an unusual, unaccepted, illegal, or harmful behavior towards the society. Deviance on social media can include deviant groups, deviant acts, deviant events, or deviant tactics.

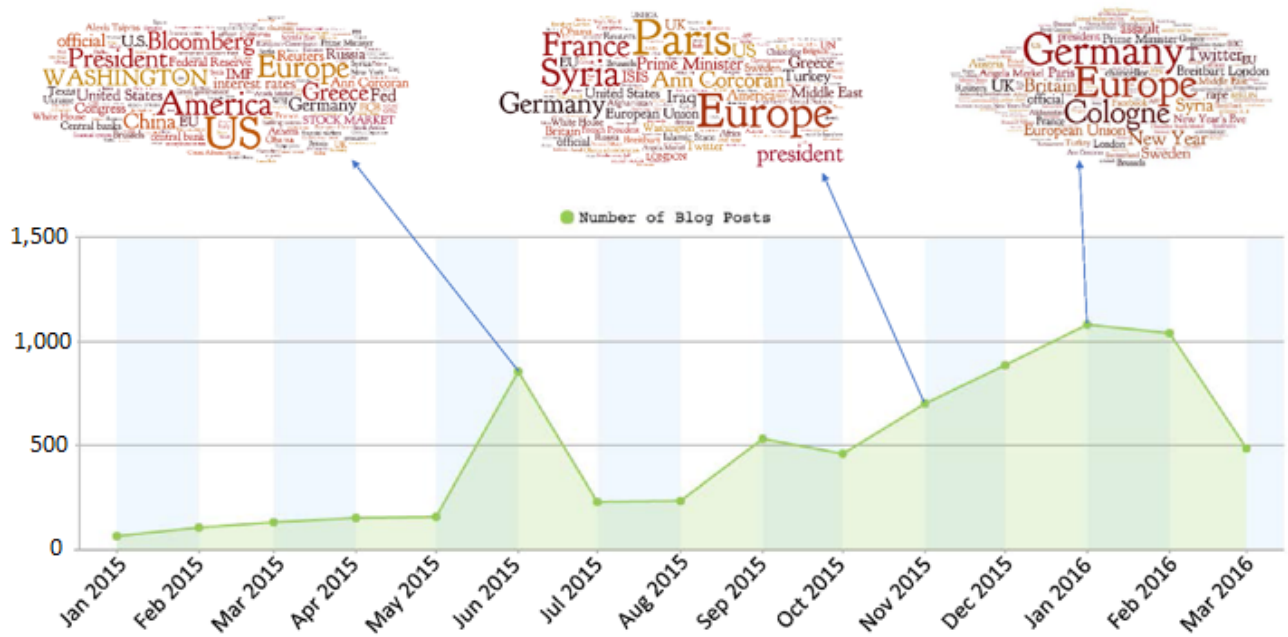


Figure 1: Posting trend during migrant issue in European Union.

Inquiry and Word Count (LIWC) [TP10]. We also used AlchemyAPI service (available at www.alchemyapi.com) to identify the language of each blog post. We found that English was the dominant language in the collected dataset. In addition to identifying languages using AlchemyAPI service, we used it to extract 169,846 named-entities. We extracted 37,611 outbound URLs (22,298 unique URLs) from 3,376 different domains. We found a few of highly linked domains/websites subscribe to extreme right-wing ideologies and supported highly isolationist and anti-immigrant views (e.g., Breitbart, ZeroHedge).

Finally, we loaded our cleaned and enriched data into the Blogtrackers [Blo18] database for analysis. Since migrant issue in Europe gained maximum traction in 2015, we selected January 2015 to March 2016 as the period of our analysis. We used posting frequency feature of Blogtrackers to study posting trends and identify any unusual patterns in blog postings from January 2015 to March 2016 (Figure 1).

4 Research Findings

We observed a spike in activity started in June 2015 with a consistent increase in posting activity from November 2015 through January 2016. To study these high posting time-periods, we extracted the top named-entities to identify topics of interest for the month of June 2015 (spike in activity) and the period from November 2015 through January 2016 (months with a consistent rise in activity). For June 2015, we found the “US”, “America” and “Washington” among the top 10 entities, “Europe” was third on the list for this time-period indicating that the blogosphere was interested in the United States of America in this period. Similarly, for the period of November 2015 to January 2016, we found “Europe”, “Syria”, “France”, “Germany”, and a few European cities as the top entities for each month from November 2015 to January 2016 indicating a shift in interest from the USA to Europe.

To further study the shift in interest in the blogosphere, we selected a few top entities from the above time periods and ran *keyword trends analysis*. Keyword trends analysis is a feature of Blogtrackers that uses document-frequency analysis to provide the overall trend for keywords of interest. We ran keyword trends analysis for keywords “America”, “China”, and “Europe” (Figure 2) and a separate keyword trends analysis for keywords “Europe”, “Germany”, “Paris”, “Syria”, and “migrant” (Figure 3) for the period January 2015 to March 2016. Using the *keyword trends analysis* for “America”, “China” and “Europe” (Figure 2), we found that America was the primary focus in the blogosphere in June 2015 but it eventually declined and the primary focus changed to Europe. Upon further investigation, we found that the US was involved in multiple sub-events in June 2015. Some of the noteworthy sub-events were - 815 Syrian refugees were admitted into US and was admitting

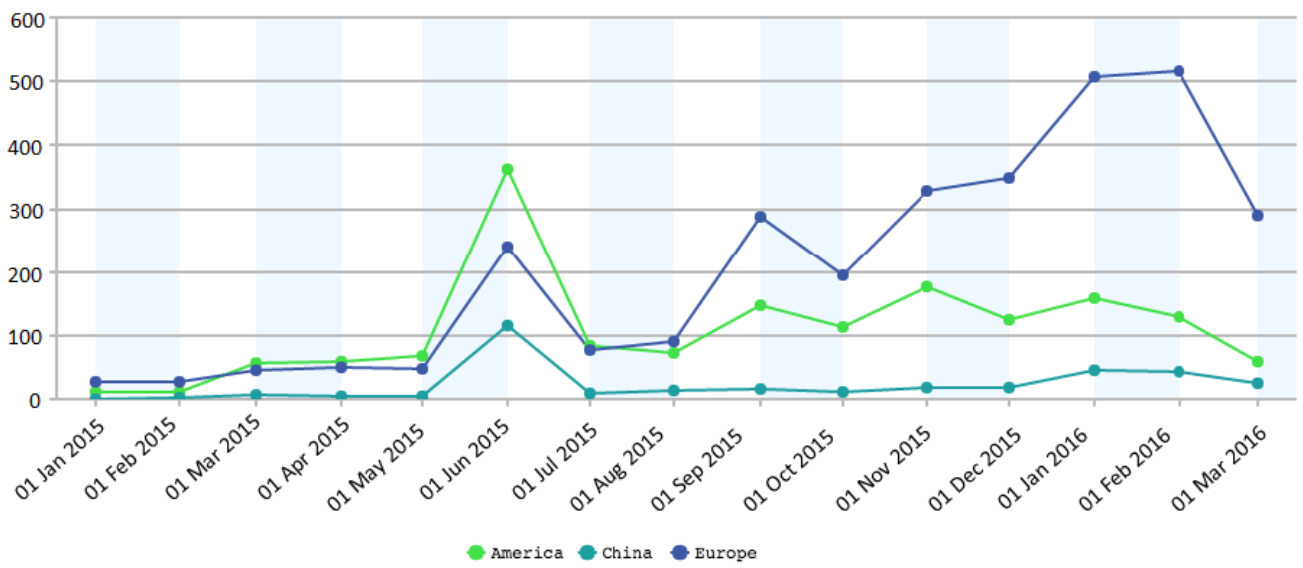


Figure 2: Keyword trends “America”, “China”, and “Europe”.

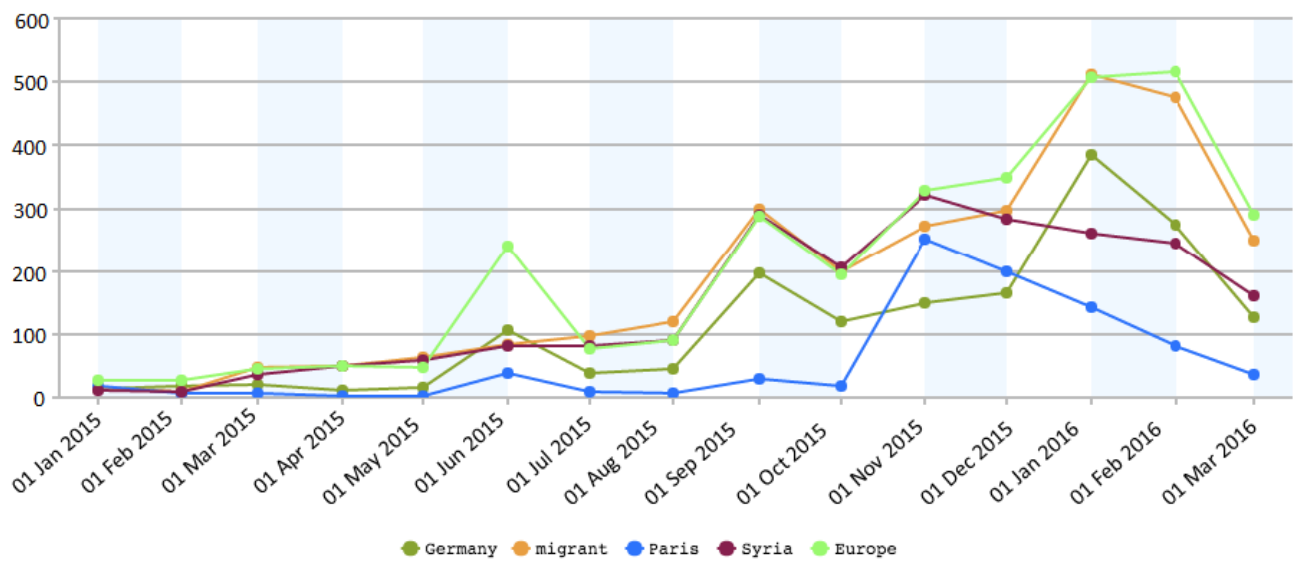


Figure 3: Keyword trends “Germany”, “migrant”, “Paris”, “Syria”, and “Europe”.

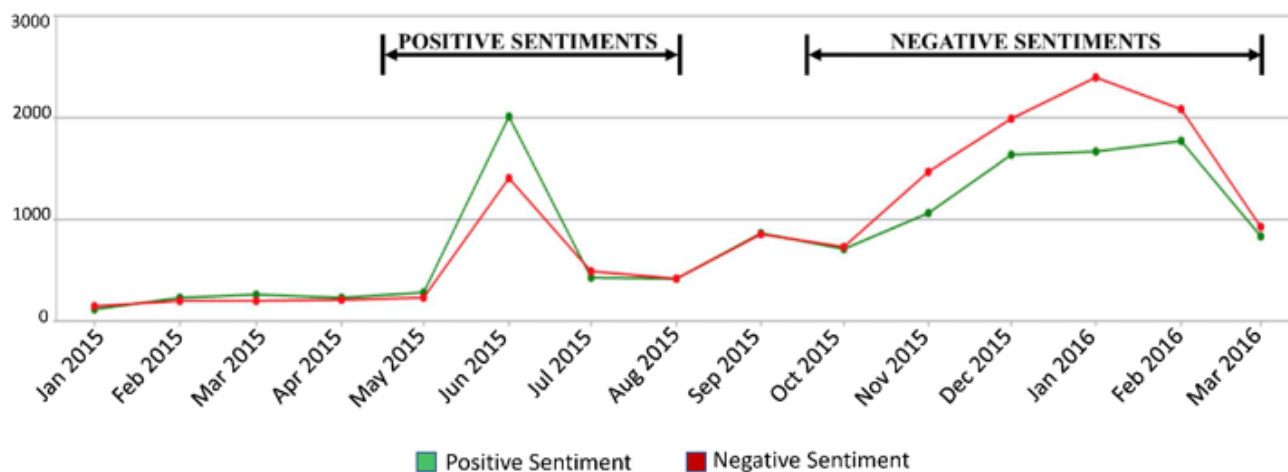


Figure 4: Sentiment trends.

11,000 more [noa15], US participation in fighting ISIS [noae], rise in tensions between US and China [noab, noaf], Russias travel ban on European and American elites [noaa], and FIFA officials investigated by FBI on corruption charges [noad].

Using the keyword trend analysis for “Europe”, “Germany”, “Paris”, “Syria”, and “migrant” (Figure 3) we observed the trend for the keywords “Europe” and “migrant” was almost identical indicating that these blogs were very relevant to the migrant crisis in Europe. We also observed a sudden rise for the keyword “Paris” around Paris attacks in November 2015 [noaa]. In addition to that, we observed a sudden rise in the keyword “Germany” during January 2016. Upon further investigation, we found that blogs [noac, Med16, J16] were discussing the alleged reports of German women being harassed, assaulted, and raped by refugees during the 2016 New Years Eve celebrations and a massive outcry against the media and government for improper handling of the situation [Ww, Sma16].

To further analyze the sentiments reflected in the blogs, we conducted *sentiment trend analysis* using Blogtrackers for the period January 2015 to March 2016. Figure 4 shows the sentiment trends of this period of time. The overall sentiment was mostly positive from May 2015 to July 2015, neutral from January 2015 to May 2015, and July 2015 to October 2015, but after October 2015 there was a flip in the sentiments from positive to negative. To investigate deeper and understand the narratives toward migrants/refugees, we identified all the sentences containing words “migrant” or “refugee” and extracted sentiments. Figure 5 shows average sentiment trends toward migrants/refugees, following a similar pattern of a flip from positive to negative. Upon investigating, we found that the overall narrative in mainstream media toward migrants was positive. Initially, citizens of many European countries sympathized and wanted their government to help the refugees. People even raised “Refugees Welcome” banners at major soccer events (Figure 6). However, their sentiments had changed after October 2015 which might be attributed to the Paris attacks in November 2015 and assaults on German women in January 2016. People were rattled by these events and protested by raising “Rapefugees not Welcome” banners (Figure 6).

5 Conclusion and Future Work

Social media has evolved from merely being a hub for high school students used socialize with friends to a platform that can be used for agenda-setting. This behavior of content framing encourages the development of weaponized narratives that can influence readers toward deviant acts or disorient them in an attempt to provoke hysteria. In this paper, we used the migrant crisis in EU as a case study to observe the change in sentiments of citizens toward migrants and to understand the shift in a narrative on the blogosphere. We explained the use of targeted sentiments to study any shift in narrative towards any entity (in the case above - migrant/refugee).

We followed a similar approach - using social media posts’ date for analysis and anchor all events being discussed in the social media post to its publication date - that others used to conduct research on social media which introduce some limitations. In future, we plan to identify any temporal expression from the post itself to

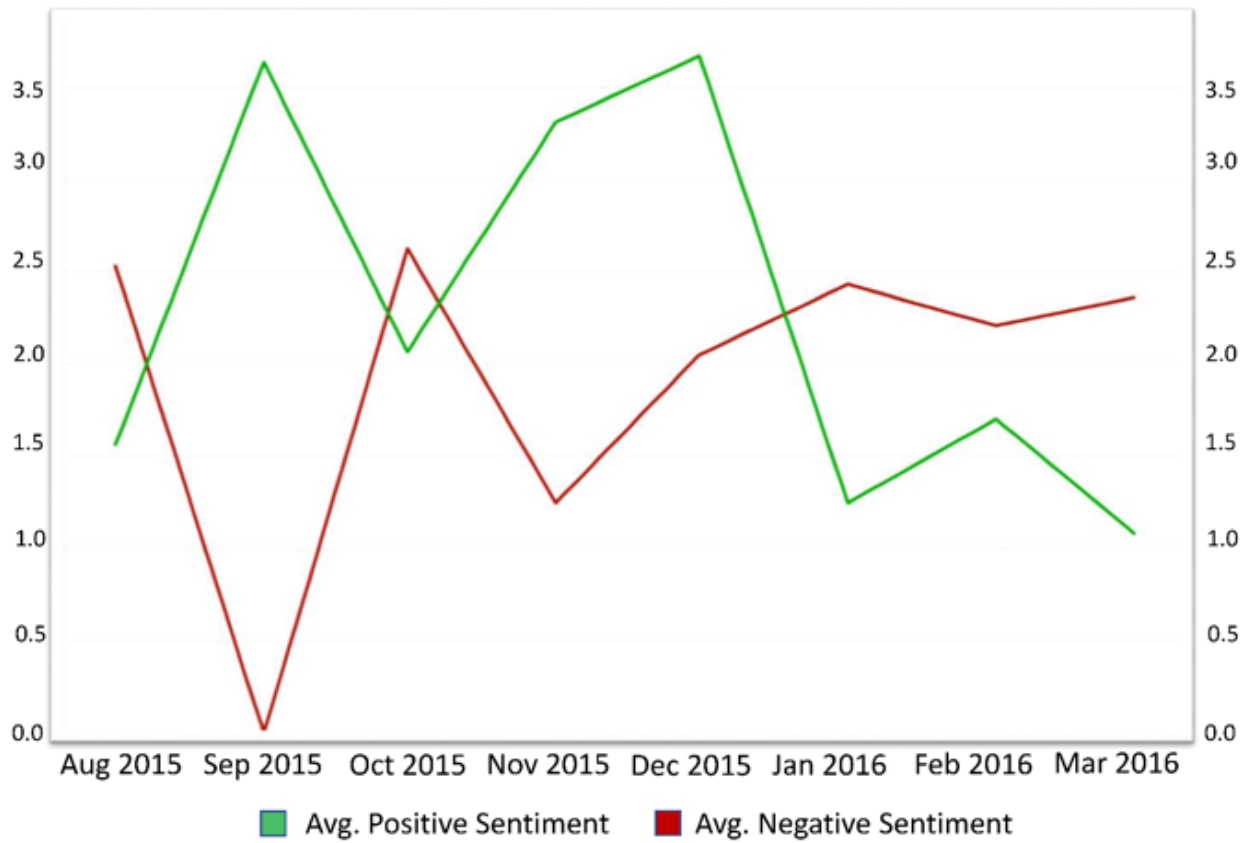


Figure 5: Targeted sentiments toward migrants/refugees.



Figure 6: Banners during migrant issue in European Union.

have more precise analysis.

In this research, we primarily focused on studying the shift in narratives. For future, we plan to build a model to identify narrative for any given text. Employing such a model to continuously monitor streaming social media content can help detect any deviant narratives like fake news or propaganda. Being able to identify deviant narrative at an early stage can help stem the spread of deviance on social media and build effective counter-narratives.

Acknowledgements

This research is funded in part by the U.S. National Science Foundation (IIS-1636933, ACI-1429160, and IIS-1110868), U.S. Office of Naval Research (N00014-10-1-0091, N00014-14-1-0489, N00014-15-P-1187, N00014-16-1-2016, N00014-16-1-2412, N00014-17-1-2605, N00014-17-1-2675), U.S. Air Force Research Lab, U.S. Army Research Office (W911NF-16-1-0189), U.S. Defense Advanced Research Projects Agency (W31P4Q-17-C-0059) and the Jerry L. Maulden/Entergy Fund at the University of Arkansas at Little Rock. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the funding organizations. The researchers gratefully acknowledge the support.

References

- [Alb16] Jonathan Albright. The #Election2016 Micro-Propaganda Machine, November 2016.
- [Blo18] Blogtrackers. Welcome to Blogtracker, 2018.
- [CHFA11] Wen-Ying Sylvia Chou, Yvonne Hunt, Anna Folkers, and Erik Augustson. Cancer survivorship in the age of YouTube and social media: a narrative analysis. *Journal of medical Internet research*, 13(1), 2011.
- [CRF12] S. Corman, Scott W. Ruston, and Megan Fisk. A pragmatic framework for studying extremists use of cultural narrative. In *2nd International Conference on Cross-Cultural Decision Making: Focus 2012*, pages 21–25, 2012.
- [Gri17] Elizabeth Grieco. More Americans are turning to multiple social media sites for news, November 2017.
- [HOB⁺17] Muhammad Nihal Hussain, Adewale Obadimu, Kiran Kumar Bandeli, Nooman Mohammad, Samer Al-khateeb, and Nitin Agarwal. A Framework for Blog Data Collection: Challenges and Opportunities, June 2017.
- [J16] Oliver JJ Lane5 Jan 20164 and 428. Cologne Mayor: Women Should Be More Careful After Migrant Mass Rapes, Promises 'Guidance' So They Can 'Prepare', January 2016.
- [Med16] Fred Alan Medforth. Allah's Willing Executioners: Muslim mass sexual assaults in Germany on New Year's Eve, January 2016.
- [MHN⁺] Esther Ledelle Mead, Muhammad Nihal Hussain, Mohammad Nooman, Samer Al-khateeb, and Nitin Agarwal. Assessing Situation Awareness through Blogosphere: A Case Study on Venezuelan Socio-Political Crisis and the Migrant Influx.
- [Mir] Holmstrom Miranda. Miranda Holmstrom. The narrative and social media. | StratCom.
- [noaa] Attacks in Paris. *The New York Times*.
- [noab] China Blamed For "Largest Theft Of US Government Data Ever" - 95% Of Federal Employees Affected | Zero Hedge.
- [noac] The "Monsters" Unmasked: Cologne Police Admit "Most Of The Attackers Were Refugees".
- [noad] The Next Escalation: FBI Launches Probe Of Russia 2018 World Cup Award.
- [noae] US Will Send 400 More Troops To Iraq Bringing Total To 3,500; Open New Military Base.

- [noaf] Vietnam To Buy Fighter Jets, Drones From West To Counter Chinese 'Aggression' | Zero Hedge.
- [noa15] To date, 815 Syrian refugees admitted to US; 92% are Muslim, May 2015.
- [noa18] So What Did We Learn? Looking Back on Four Years of Russias Cyber-Enabled Active Measures, January 2018.
- [RCS+16] Scott W. Ruston, J. V. Cohn, S. Schatz, H. Freeman, and D. J. Y. Combs. More than just a story: Narrative insights into comprehension ideology and decision-making. *Modeling Sociocultural Influences on Decision Making: Understanding Conflict, Enabling Stability*, page 27, 2016.
- [Rie93] Catherine Kohler Riessman. *Narrative analysis*, volume 30. Sage, 1993.
- [SG17] Elisa Shearer and Jeffrey Gottfried. News Use Across Social Media Platforms 2017, September 2017.
- [Sma16] Alison Smale. As Germany Welcomes Migrants, Sexual Attacks in Cologne Point to a New Reality. *The New York Times*, January 2016.
- [TP10] Yla R Tausczik and James W Pennebaker. The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of language and social psychology*, 29(1):24–54, 2010.
- [Ww] Deutsche Welle (www.dw.com). Report: Cologne New Year's Eve attacks 'could have been prevented' | News | DW | 17.03.2017.