Differences in Emotional Reactions to Social Media Content: The Role of Location and Content Type

Farhad Mohammad Afzali
University of Nebraska at Omaha
Omaha, NE, USA
farhadafzali@fulbrightmail.org

Christian Haas
University of Nebraska at Omaha
Omaha, NE, USA
christianhaas@unomaha.edu

Margeret Hall
University of Nebraska at Omaha
Omaha, NE, USA
mahall@unomaha.edu

ABSTRACT

Individuals in diaspora tend to engage and interact with the people of their origin. Afghans in the 1980s and 90s, still one of the largest diaspora communities, used traditional channels such as Radio and Television to communicate with their loved ones inside Afghanistan. The Internet facilitated changing their mode of communication to near real-time updates. Diaspora and local communities engage in the same Social Awareness Streams (SAS), but there is a lack of research focusing on difference and similarities of these communities. This paper investigates the differences in emotional disclosure and expression between diaspora and local communities in Afghanistan using data from 'Afghanistan My Passion', the largest public Afghan Facebook community page. We investigate a corpus of 2,165 Persian language words considering the location and gender of the social media user. This work provides the first analysis towards understanding differences in diaspora and local communities' emotional reactions to social media content and extends the body of literature on Persian language sentiment analysis.

CCS CONCEPTS

• **Human-centered computing** → **Social media**; **Social networking sites**.

KEYWORDS

social networks, social awareness streams, sentiment analysis

ACM Reference Format:
Farhad Mohammad Afzali, Christian Haas, and Margeret Hall. 2019. Differences in Emotional Reactions to Social Media Content: The Role of Location and Content Type. In *11th ACM Conference on Web Science (WebSci '19), June 30-July 3, 2019, Boston, MA, USA.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3292522.3326035

1 INTRODUCTION

Diaspora communities are necessarily linked with displacement, dispersal, and migration [21], yet researchers have different opinions regarding its characteristics. Conner considers individuals living outside their homeland as diaspora communities [8], while another study describes diaspora as migrants who are immigrants, guest workers, asylum seekers, ethnic minorities, and displaced populations as [21]. Hiller, adding to the definition of diaspora, states that there are three phases of migration cycle: pre-migrants, post-migrants, and settled migrants [11]. Afghan refugees make a large portion of the world's total diaspora community. We could not find Afghan Diaspora statistics around the world, but following refugee statistics provide a sense of the size of the Afghan diaspora. Afghan refugees had the largest worldwide refugee community for 32 years [19] before being surpassed by Syrian refugees in 2014 [1]. Afghans started fleeing the country in 1978 – after the Saur revolution – and outward migration has continued since then. In 2014, there were over 3.7 million Afghan refugees, with most of them living in Iran and Pakistan [19]. In comparison, Afghanistan has a population of over 34 million people as of 2017.

Afghans in diaspora have had an active role in the country's politics, reconstruction, and capacity building. Since 2001, Afghan presidents – Hamid Karzai and Ashraf Ghani – have been diasporic individuals [23]. Furthermore, three of four political groups attending the 2001 Bonn talks – gathered to decide the transitional government of Afghanistan after 9/11 – came from the diaspora [22]. Besides the country's leadership, diaspora members were given key positions such as ministries in the interim governments [23]. Nassery's report on the role of the Afghan-American diaspora in peacebuilding and reconstruction states: "Individual and group behaviors are an important element as of peacebuilding as are education or skill level, and by behaving professionally and collectively, the Afghan-American diaspora can best influence policy planning and implementation of reconstruction in Afghanistan" [[18], p. 1].

Diaspora communities – Afghans in the above mentioned cases – play a significant role their origin homeland. However, there is limited research on how Afghan diasporic identities are shaped and reshaped as they interact with their homeland. [9] discusses that 21st century is quite different from prior diasopras as new technologies pave the way for a faster and real-time communication to virtually engage in the events happening in their homeland. Social media's broad and deep discourses allow researchers to observe these interactions unobtrusively [14]. Social media has affected the integration and presence of diaspora communities as they virtually engage with their homeland at the expense of physical engagement with their adopted communities [13], which helps with studying the mutual impact of locals and diaspora alike. Afghans inside and outside of their country discuss issues and share opinions on current affairs using Social Awareness Streams (SAS) – a kind of semi-public communication used in social media. However, thoughts and emotions regarding events could differ between people based on their location and gender. An Afghan living in diaspora might

WebSci '19, June 30-July 3, 2019, Boston, MA, USA

ACM ISBN 978-1-4503-6202-3/19/06…$15.00
https://doi.org/10.1145/3292522.3326035

Table 1: Afghanistan My Passion fans by country

#	Country of Residence	#Fans
1	Afghanistan	1,592,397
2	Pakistan	475,146
3	Germany	142,278
4	Saudi Arabia	128,189
5	Iran	98,723
6	India	89,529
7	United Arab Emirates	88,806
8	Turkey	70,319
9	United States of America	50, 074
10	Bangladesh	43,922

see corruption in the public sector differently compared to an individual residing inside the country and being part of the system. Similarly, a woman might view a new law introduced by the government differently compared to how a man sees it. For analyses of discourse in countries with high external populations, correct pre-classification is a necessary step. However diasporic communication and its differences compared to local populations is a mostly unaddressed subject in social media analysis. Also, it is unknown if different content types – posted by moderators – activate higher or lower emotional responses. This leads to our research questions:

RQ1: How do location and gender impact the sentiment of users' emotional expression?

RQ2: Which types of posts generate high emotional responses?

We chose Facebook as our data source to investigate these research questions as it is the most used social network in Afghanistan [1] and is the most accessible communication service in the country. As of 2015, mobile service providers, covering 89% of the population [2], offered Facebook packages where the subscriber did not need to have an active mobile data or Internet subscription to use Facebook [2]. It cost around one U.S. dollar to activate this service for one month. The coverage has increased and rates has gone down since then. 'Afghanistan My Passion' is the Facebook page with the largest audience among Afghans inside the country and in diaspora. The page – launched in 2011 – had 3.9 million fans in 2018, 42.9% of whom live inside the country [3] (see Table 1 for the top ten countries with the highest number of fans). The page admin posts various types of events happening across the country.

2 RELATED WORK

2.1 Diasporic Communication

Tsagarousianou states that connectivity defines diaspora better than displacement [21]. Diaspora communities have been communicating with their loved ones in their first home over generations. The type of communication has changed over time though. If we think about two centuries ago, it would take months for someone to communicate a single message with people in their homeland, probably by sending a letter and receiving a response. The emergence of technology and the facility to communicate with others without a binding to distance or time has changed the shape of diasporic communications. The Internet brings members of a community closer to a global village where all the members have real-time connectivity to any other member of the community [16]. Media technology like social media, voice over internet protocols (VoIP) and diasporic media have a crucial role in forming diasporic identities and diaspora in general [21]. With media technology facilitating interactions among the diaspora communities, the concept of e-diaspora is getting popular [16]. Research in the role of ICT (Information and Communication Technology) in the lives of diasporic peoples contributes to a definition of the concept of e-diaspora, understood here as the result of the unwilling dispersal of members of a nation across several countries [16].

2.2 Social Network Sites (SNS)

Social network sites or social networking sites are web-based services used by individuals for constructing a public or semi-public profile within a bounded system, forming a list of users with whom they may connect, and exploring the list of connections [4]. Users interact online by creating profiles and providing (semi)personal information in the form of text, photos and other media [10]. As social networking and media platforms are generally based on true identities or variants thereof [10, 15], they are well suited for digital community analyses [10]. Social media and social networking sites have quickly ascended from a novelty of the early 2000's to a fact of life, and daily necessity. SNS like Facebook, LinkedIn, Twitter, Google Plus have become extremely popular in the recent years [5]. With the help of SNS, users – diasporic individuals – create social identities by creating social profiles and building social networks, which creates a sense of belonging to the network. Users consider themselves as part of the community psychologically if they: see themselves emotionally involved with the online community; evaluate self-worth based on the membership to a particular social group; or sense self-awareness of being part of the group [6].

2.3 Social Awareness Streams (SAS)

Social Network Sites have changed the shape of communications and interactions among people. SAS is one of the types of communication services that leverage the SNS. A kind of semi-public communication, i.e., Facebook and Twitter, which are very popular [12]. In this communication style, a user posts messages, photos or videos. What a user posts will appear to his/her contacts' news feed. The users can react synchronously or asynchronously with, for example, a like, comment or share, to what they see in their news feed posted by other people in his/her network [17]. Naaman et al. argue that the public nature of communication, the brevity of published content, and a highly connected social space are the three main factors that distinguish SAS from other types of communication services [17]. Therefore, we chose a Facebook page, Afghanistan My Passion, for this study.

3 RESEARCH DESIGN

In order to answer our research questions, we leveraged Python scripts to extract the data from Facebook. Between 2011-2017, there were 8,404 status updates by page admins. The total number of

[1] https://www.internews.org/news/social-media-afghanistan-users -and-engagement
[2] https://www.independent.co.uk/life-style/gadgets-and-tech/features/how-social-media-is-empowering-young-afghan-women-the-facebook-effect-10375022.html
[3] https://www.socialbakers.com/statistics/facebook/ pages/total/afghanistan

comments on these mentioned statuses is 646,399. We focused on data between July 13, 2017, and August 07, 2017, specifically on comments with 150+ words (n=480) because all types of topics found across all the page's posts were covered in that three week interval. We included only text written in Persian (Dari dialect) using Persian characters. English, transliterated texts, and other languages are excluded from this initial study as current language processing packages are unable to simultaneously analyze multiple languages or indeed even interpret the same language written in differing scripts. 91 of the total comments in our data set were English, Pashto, or transliterated. We removed those comments to keep our analysis based on Persian (Dari). After cleaning the comments, we had 389 of them to extract our sample from. Due to lack of availability of automatic processing tools for Persian data classification and manual identification of comment author's gender and location, we took a sample of 100 comments as to manually identify the gender and location of comment authors [4]. We used Quintly [5] to get the demographics of the fans based on the country they reside in. Quintly is a web-based social analytics tool that helps with scraping of information from social media sites.

3.1 Data Classification

To classify the data in common themes of content types, four Persian native speakers analyzed and coded the data individually. The coders used a jury format to reach a consensus regarding the common themes. The common themes (post types) identified were Cultural, Religious, Historical, Education, Racial Discrimination, Security Situation, and Healthcare. We then identified the comment author location and their gender by viewing their Facebook profiles. We dropped three profiles for which locations were not identifiable and two profiles for whose gender could not be identified.

3.2 Sentiment Analysis Tools

Following the data collection and classification, the biggest challenge was the lack of available tools that can analyze Persian/Dari sentiment. SentiStrength [20] was one of the tools that support Persian sentiment analysis. However, we found out that their Persian dictionary is still not mature as it leaves most of the words uncategorized. We used Polyglot [3] as the tool for sentiment analysis as it categorized most of the words, as it proved to be the most efficient sentiment analysis tool for the Persian Language. Polyglot is a python package that has robust features such as tokenization, language detection, Named Entity Recognition, Parts of Speech Tagging, Sentiment Analysis, Word Embedding, Morphological Analysis, and Transliteration of 100+ languages. It determines the polarity of words based on their negativity and positivity by assigning scores of -1 (negative words), 0 (neutral words), and +1 (positive words) to the words respectively. The analysis found 2,165 emotionally intoned terms in the posts and comments.

4 ANALYSIS AND RESULTS

In this section, we describe the analyses we performed on the data and the corresponding results. First, a word frequency-based analysis was conducted to determine commonalities between admins and fans, as well as to identify popular topics. Second, a sentiment analysis was performed to evaluate potential differences in sentiment across different topics, groups, and genders.

[4]https://www.socialbakers.com/statistics/facebook/ pages/total/afghanistan/
[5]https://www.quintly.com

Figure 1: Frequency of words in posts by admin.

Figure 2: Frequency of words used in comments by fans.

4.1 Word Frequencies and Topics

As a first analysis, we counted the frequency of the words used in all posts and comments since the start of Afghanistan My Passion, 2011. It helped us understand what the most common topics were and how the fans reacted to the posts. We generated two word clouds using wordclouds.com (see Figure 1 and 2). The figures demonstrate that there was a similarity in the topics discussed in posts by the admin and comments shared by the users. Additionally, the major issues addressed were positive, which indicates that the fans regardless of their location have similar sentiment about the topics discussed.

The word cloud in Figure 1 demonstrates the words used in the posts made by the admin of the page. The five most frequent words used by admins in the posts (Figure 1), from highest to lowest, were افغانستان، جان، زن، سال، وطن. English translations of these words are Afghanistan, life (or dear), woman, year, and homeland. Besides the five terms mentioned, most of the high frequency used words had positive messages in them. The words used in comments by fans (Figure 2) had positive tones in them as well. The most frequently used word used in comments was identical to the one used in posts, افغانستان. The remaining four highest frequency

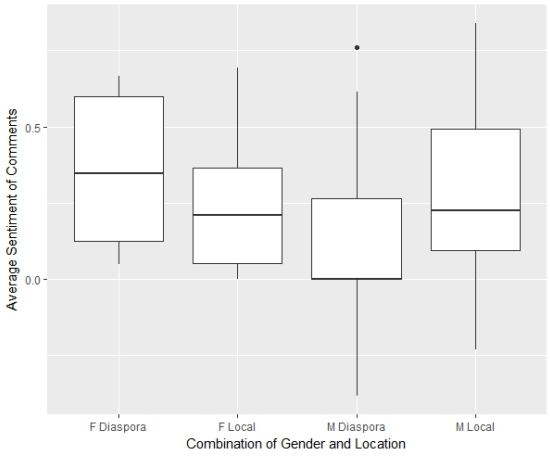

Figure 3: Comment sentiment based on gender and location

used words were زنده، وطن، جان، خداوند. English translation of these words are (highest to lowest): Afghanistan, God, life (it means dear, too), homeland, and alive. Looking at the most frequent words, frequently discussed topics tend to be related to religion, life, and security. Combining the top ten words used in the statuses and comments we found that they all had positive valence. While more than a half of the fans reside outside the country, the most frequent words were about 'Afghanistan' and 'homeland'. This indicates that Afghans in diaspora discuss the country and the situations in the country as often as the local Afghans do, revealing an active interest that the diaspora has for its home country.

4.2 Sentiment Analysis

For the subsequent analysis, we concentrated on the comment-level sentiment. Comment sentiment was defined as the average percentage of positive and negative words in a comment; thus ranging from -1 to 1. For example, if a comment had two positive and one negative word, the average sentiment score was (2 - 1)/3 = 0.33. Figure 3 shows the distribution of comment sentiment based on location and gender. The average comment sentiment is 0.309 for Local and Male, 0.248 for Local and Female, 0.162 for Diaspora and Male, and 0.339 for Diaspora and Female. We had a total of 14.5% comments from women, and 43.5% of the comments were from Diaspora. An analysis of the total number of positive and negative words per comment revealed no statistically significant difference based on Location or Gender. To address the RQs, we conducted several statistical analyses based on the previously described data. Analysis of Variance (ANOVA) was used for the subsequent analyses, as the assumption of Gaussian distributions of the sentiment scores cannot be rejected at the 0.05 level. First, we considered the effect of gender (M/F) and location (Local/Diaspora) on the average comment sentiment. A two-way ANOVA with interaction effect found that neither main or interaction effects were significant at the 0.05 level, however the p-value of the location impacting the average comment sentiment was close to significance with 0.055 (see Table 2). Subsequent t-tests for the subgroups revealed the difference between Local and Diaspora was significant for Male comments based on a p-value of 0.02. Neither gender nor the interaction of gender and location was significant. Next, we included the type of post in the analyses. Post types were either cultural, religious, educational, historical, healthcare, racial discrimination, and security. Including the post type in the ANOVA did not change the resulting outcome, i.e., neither the post type nor its interactions with gender/location were found to be significant (see Table 3). This can, however, be a result of the currently used number of comments and could change by including a larger number of comments in subsequent studies. Third, when using an ANOVA with only the post type groups, we found that the post type significantly influences the number of positive/negative words in the comments at the 0.05 level (see Table 4). Specifically, Fisher's Least Significant Difference test showed that posts concerning Racial Discrimination lead to a higher number of positive/negative sentiment words than posts about culture, religion, security, or education. Also, history related posts lead to a higher number of words than posts about religious topics. This indicates that Racial Discrimination in particular seems to lead to more discussions on Facebook.

Table 2: Effects of gender and location on sentiment score

	Df	Sum Sq	Mean Sq	F	p-value
Gender	1	0.076	0.07629	1.042	0.311
Location	1	0.279	0.27941	3.817	0.055
Gender:Location	1	0.145	0.14488	1.979	0.164
Residuals	65	4.758	0.07320		

Table 3: Effects of gender, location, and post type on sentiment score

	Df	Sum Sq	Mean Sq	F	p-value
Gender	1	0.076	0.07629	1.002	0.322
Location	1	0.279	0.27941	3.670	0.061
Post_Type	6	0.282	0.04696	0.617	0.716
Gender:Location	1	0.193	0.19297	2.534	0.118
Gender:Post_Type	4	0.441	0.11023	1.448	0.232
Location:Post_Type	4	0.104	0.02606	0.342	0.848
Residuals	51	3.883	0.07614		

Table 4: Effects of post type on positive/negative words

	Df	Sum Sq	Mean Sq	F	p-value
Post_Type	6	2937	489.6	2.651	0.0236
Residuals	62	11451	184.7		

Similarly, Table 5 shows an overview of average sentiment scores based on the Location (Diaspora vs Local) and Post Type. The NA values indicate that no posts were observed in this category. Based on the table, Cultural topics seem to correspond to more positive sentiment as compared to other topics. In addition, differences between the Local and Diaspora sentiment are observable in posts about Racial Discrimination, Religion, and Education. However,

another two-way ANOVA with Post Type and Location as well as their interaction reveals no significance at the 0.05 level, with Location having the smallest p-value of 0.056. Finally, investigated whether there is a correlation between the sentiment of the original post (i.e., post by the site admin) and the sentiment of the comments. With a correlation coefficient of 0.08, the analysis revealed that there is little correlation between these two aspects.

Table 5: Location and post type sentiment differences

Post_Type	Diaspora	Local
Cultural	0.38396624	0.2763158
Educational	0.16666667	0.3150685
Healthcare	0.33333333	NA
Historical	0.12068966	0.2222222
Racial Discrimination	0.01075269	0.5121951
Religious	0.05405405	0.2719101
Security	0.06666667	NA

5 DISCUSSION & CONCLUSION

This study explored the sentiment of diasporic individuals in comparison to locals. It also examined the interactions of communities on social media and analyzed emotional characteristics based on their location and gender. Regarding RQ1, we find no statistical differences between the populations based on their declared gender. Differences in location between locals and diasporic individuals tends towards significance, which is the basis of any future work in this direction. Our analysis demonstrated that Afghans are more sensitive to racial discrimination and engage in talking about history more than any other category. These types of posts generate the highest emotional responses (RQ2). Finally, this paper contributes to the community of practice. The social media community activists, specifically Afghan communities, can predict the sensitivity and commonality of a post before publishing it. By doing so, they can post the right content at the right time. Social media can also be leveraged for various purposes to address local and diaspora communities. Cogburn & Espinoza-Vasquez, for instance, found that leveraging Web 2.0 helped Obama 08 campaign "translate online activity to on-the-ground activity" [[7], p. 200]. Hence, we believe that activists can take advantage of the existence of social media to develop a unified virtual local and diaspora community to promote causes and/or spread awareness.

6 LIMITATIONS & FUTURE WORK

This study has several limitations. We tried to identify the sentiment of fans based on their gender and location generally. This is manually intensive and results in a small-n phenomenon. In the next phase we will examine the more extensive dataset – collecting data from several pages – with more categories of admin posts to ensure maximum coverage of the page content in our analysis. We also plan to analyze the sentiment on the user level. The current research is hampered by the lack of robust Persian language tools. We argue that introducing or identifying a sentiment analysis tool with a more complete dictionary and the sentence-level sentiment analysis functionality will be a significant contribution to the computational analysis of the Persian language. The small range of Polyglot may contribute to the lack of significance in the results, where a more fine-tuned tool like LIWC [6] may display differences across characteristics in language use.

REFERENCES

[1] 2014. Syria tops the world as source of refugees | News | Al Jazeera. https://bit.ly/1dxhXKR
[2] 2015. Ministry of Communication and IT (MCIT). https://mcit.gov.af/en
[3] Rami Al-Rfou, Bryan Perozzi, and Steven Skiena. 2013. Polyglot: Distributed Word Representations for Multilingual NLP. (2013). https://doi.org/10.1007/s10479-011-0841-3
[4] Danah Boyd and Nicole Ellison. 2010. Social network sites: definition, history, and scholarship. *IEEE Engineering Management Review* 3, 38 (2010), 16–31.
[5] Mustapha Cheikh-Ammar and Henri Barki. 2016. The Influence of Social Presence, Social Exchange and Feedback Features on SNS Continuous Use:. *Journal of Organizational and End User Computing* 28, 2 (2016), 33–52. https://doi.org/10.4018/JOEUC.2016040103
[6] Christy M K Cheung, Pui Yee Chiu, and Matthew K O Lee. 2011. Online social networks: Why do students use facebook? *Computers in Human Behavior* 27, 4 (2011), 1337–1343. https://doi.org/10.1016/j.chb.2010.07.028
[7] Derrick L. Cogburn and Fatima K. Espinoza-Vasquez. 2011. From networked nominee to networked nation: Examining the impact of web 2.0 and social media on political participation and civic engagement in the 2008 obama campaign. *Journal of Political Marketing* 10, 1-2 (2011), 189–213. https://doi.org/10.1080/15377857.2011.540224
[8] Walker Conner. 1986. The impact of homelands upon diasporas.
[9] Shehina Fazal and Roza Tsagarousianou. 2002. Diasporic communication: Transnational cultural practices and communicative spaces. *Javnost* 9, 1 (2002), 5–18. https://doi.org/10.1080/13183222.2002.11008790
[10] Margeret Hall and Simon Caton. 2017. Am I Who I Say I am ? Unobtrusive Self-Representation and Personality Recognition on Facebook. *PLOS One* 12, 9 (2017), e0184417. https://doi.org/10.1371/journal.pone.0184417
[11] Harry H. Hiller and Tara M. Franz. 2004. New ties, old ties and lost ties: the use of the internet in diaspora. *New Media & Society* 6, 6 (12 2004), 731–752. https://doi.org/10.1177/146144804044327
[12] Funda Kivran-swaine and Mor Naaman. 2009. Network Properties and Social Sharing of Emotions in Social Awareness Streams. *Media* July (2009), 379–382. https://doi.org/10.1145/1958824.1958882
[13] Lee Komito and Jessica Bates. 2009. Virtually local: social media and community among Polish nationals in Dublin. *Aslib Proceedings* 61, 3 (2009), 232–244. https://doi.org/10.1108/00012530910959790
[14] Andreas Lindner, Margeret Hall, Claudia Niemeyer, and Simon Caton. 2015. BeWell: A Sentiment Aggregator for Proactive Community Management. In *CHI'15 Extended Abstracts*, Vol. 18. ACM Press, Seoul, 1055–1060. https://doi.org/10.1145/2702613.2732787
[15] Jessa Lingel, Mor Naaman, and danah boyd. 2014. City, self, network: transnational migrants and online identity work. In *CSCW'14*. 1502–1510. https://doi.org/10.1145/2531602.25311693
[16] Marta Marcheva. 2011. The Networked Diaspora : Bulgarian Migrants on Facebook Facebook. 14, 2 (2011), 17–19.
[17] M. Naaman, J. Boase, and C.H. Lai. 2010. Is it Really About Me? Message Content in Social Awareness Streams. *Dot-Me.of-Cour.Se* (2010), 0–3. https://doi.org/10.1145/1718918.1718953
[18] Homira G. Nassery. 2003. *The Reverse Brain Drain: Afghan-American Diaspora in Post-Conflict Peacebuilding and Reconstruction.* Technical Report.
[19] Catherine Putz. 2015. Afghanistan's 32-Year Refugee Crisis. http://thediplomat.com/2015/09/afghanistans-32-year-refugee-crisis/
[20] Mike Thelwall, Kevan Buckley, and Georgios Paltoglou. 2012. Sentiment strength detection for the social web. *Journal of the American Society for Information Science and Technology* 63, 1 (1 2012), 163–173. https://doi.org/10.1002/asi.21662
[21] R Tsagarousianou. 2004. Rethinking the concept of diaspora: mobility, connectivity, and communication in a globalised world. *Communication and Media Research Institute University of Westminster* 1, 1 (2004), 52–65. https://doi.org/10.1017/CBO9781107415324.004
[22] Nicholas Van Hear. 2014. From" Durable Solutions" to" Transnational Relations": Home and Exile Among Refugee Diasporas. *Occasional Paper* 23 (2014), 232–251.
[23] Wolfram Zunzer. 2004. Diaspora Communities and Civil Conflict Transformation. *Berghof Occasional Paper* 26 (2004), 1–56. http://hdl.handle.net/10454/4186

[6] https://liwc.wpengine.com/