Sudhahar, S., & Cristianini, N. (2018). Detecting Shifts in Public Opinion: A Big Data Study of Global News Content. In *Advances in Intelligent Data Analysis XVII : 17th International Symposium, IDA 2018, 's-Hertogenbosch, The Netherlands, October 24–26, 2018, Proceedings* (pp. 316-327). (Lecture Notes in Computer Science; Vol. 11191). Springer, Cham. https://doi.org/10.1007/978-3-030-01768-2_26

Peer reviewed version

Link to published version (if available):
10.1007/978-3-030-01768-2_26

Link to publication record in Explore Bristol Research
PDF-document

Detecting Shifts in Public Opinion: a big data study of global news content

Saatviga Sudhahar and Nello Cristianini

University of Bristol, UK
saatviga.sudhahar@bristol.ac.uk, nello.cristianini@bristol.ac.uk

Abstract. Rapid changes in public opinion have been observed in recent years about a number of issues, and some have attributed them to the emergence of a global online media sphere [1, 2]. Being able to monitor the global media sphere, for any sign of change, is an important task in politics, marketing and media analysis. Particularly interesting are sudden changes in the amount of attention and sentiment about an issue, and their temporal and geographic variations. In order to automatically monitor media content, to discover possible changes, we need to be able to access sentiment across various languages, and specifically for given entities or issues. We present a comparative study of sentiment in news content across several languages, assembling a new multilingual corpus and demonstrating that it is possible to detect variations in sentiment through machine translation. Then we apply the method on a number of real case studies, comparing changes in media coverage about Weinstein, Trump and Russia in the US, UK and some other EU countries.

Keywords: Media Content Monitoring · Public Opinion · Sentiment analysis · Machine Translation · Big data

1 Introduction

The past few years have been marked by rapid changes in public attitudes and sentiment about a range of topics. Examples include attitudes about sexual harassment, social media, and varying degrees of support about Russia. Journalists and Social scientists have been interested for a long time in detecting, tracking and measuring rapid changes in coverage of specific issues and entities [1, 2, 5, 7], a quest that is made harder by the global nature of the media system. Studies have included analysis of Twitter content and news-media content, for example following the Fukushima disaster [3, 6] and public sentiment about migration in Britain [9]. Some of these works involve human coding methods which may be more accurate but they are often limited in their ability to deal with very large data sets. This work can be automated by collecting online news, on a global scale, and analyzing their contents to extract sentiment specific to a given entity.

Big data technologies along with AI could be part of tracking and making sense of these global shifts, perhaps in real time, by monitoring global news media coverage. This poses the challenge of measuring sentiment across different languages in a comparable way. In turn, this requires a consistent method for calibrating and comparing the sentiment extracted from news in different languages.

In this paper we present one approach to detect shifts in media coverage about specific issues, in a multilingual setting. On the technical side, we isolate sentiment about a topic by extracting the words that immediately surround each mention of that topic, and analyzing them with the LIWC (Linguistic Inquiry and Word Count) dictionary [11]. This is done both in English, and in machine-generated English (machine-English for short) produced by Moses, the statistical phrase-based translation tool [10]. In order to measure the validity of this approach, we assemble a new multilingual corpus of news articles, published by Euronews in English, German, French, Spanish and Italian, covering the same topics, and we compare the sentiment about two chosen entities 'Europe' and 'Russia', extracted from the English version with that extracted from the machine-English versions, reporting significant correlation both in positive sentiment and negative sentiment (ranging between 15% and 60% depending on the language pair). We continue by demonstrating the technology in action on a larger set of news outlets from 6 countries (US, UK, Germany, Italy, France and Spain), analyzing a total of 4.4M articles, tracking the changes in positive and negative sentiment surrounding three entities: Weinstein, Trump and Russia. We observe correlations but also interesting patterns of difference, which may reveal different attitudes about the same entities. We demonstrate how this could be used to monitor rapid changes in sentiment and attention about any given entity in global news media at a large scale.

In Section 2 we discuss related work in the domain of analysing opinion shifts in online news and social media. In Section 3 we describe the data and methods used. Section 4 discusses how sentiment can be measured through machine translation. In Section 5 we discuss results related to the amount of attention and sentiment for Weinstein, Trump and Russia and in Section 6 we discuss conclusions.

2 Related Work

Studies in the past have identified public opinion shifts in social media and opinion polls. People have proposed methods for opinion mining of specific entities from Twitter content [4]. Measures of public opinion derived from polls related to consumer confidence and presidential job approval polls have seen to be correlated with sentiment measured from analysis of text from tweets [16]. This work detected topics and measured sentiment around these topics using the subjectivity lexicon from Opinion Finder [17]. Another work [18] studied how emotions, and their role in public opinion formation, can be tracked in online forums. Sentiment is measured using the ANEW (Affective Norms for English Language Words [25]) list measuring three different kinds of emotions: valence, arousal, and dominance. They showed that some political events such as the 9/11 attack unleash disagreement in valence and arousal than in dominance emotions and they have different lasting effects. Opinion shifts have been tracked in real time

[19] through the introduction of computational focus groups in Twitter. Users were grouped according to similar user biases in to average users and elites and then the response of these groups to US presidential debates were tracked and shifts in sentiment were found.

Multilingual sentiment analysis has been explored by researchers over the past few years. Previous work [21, 22] has shown that machine translation systems are mature enough to be employed to produce reliable training data for sentiment classification in languages other than English. They found that the gap in opinion and sentiment classification performance between systems trained on English and translated data is minimal. There has been a wide research effort on analyzing sentiment in news other than English by applying bilingual resources and machine translation techniques to employ the sentiment analysis approaches existing for English [24, 23]. For example sentiment of entities has been analyzed in the past from news text translated from 8 foreign language news papers in to English using information extraction methods and using the IBM Web Sphere Translation Server (WTS) to translate text to English [24]. This study was limited to data containing news articles from 10 days in May 2007 and the entities analyzed were the ones most common across news papers in their data set. Machine translation has also been used successfully to generate resources for subjectivity analysis in other languages such as Romanian and Spanish [23].

In this paper we demonstrate a large scale experiment on detecting sentiment shifts about specific issues. Our work measures sentiment using the LIWC tool, in news content across several languages, assembling a new multilingual corpus and demonstrates that it is possible to detect variations in sentiment through machine translation with Moses [10]. We apply the method on a number of real case studies and in large scale, showing that we could now monitor the global news media for rapid changes in sentiment about entities.

3 Data and Methods

The data sets used for this study were collected by our previously developed modular system [26], an integrated platform for monitoring and analyzing news media. In order to validate if sentiment across languages can be measured using machine translation, we used news articles from Euronews for French (14,646 articles), German (15,850 articles), Italian (16,886 articles) and Spanish (16,821 articles) from January 2015 to October 2017. For the rest of the analysis we obtained news articles from French, Spanish, Italian and German news outlets from January 2010 to March 2018. The most prominent news outlets for each country were selected for the analysis, which are Le Monde, Le Figaro, Libération (France), Der Tagesspiegel, Die Welt, Die Zeit (Germany), El Mundo, El País, La Vanguardia (Spain) and La República, La Stampa, Corriere della Sera (Italy). We also collected news data from UK and US outlets in the same period such as BBC, Daily Mail, Guardian, The Independent, Daily Telegraph and Daily Mirror (UK) and Seattle Times, LA Times, New York Times, Washington Post and New York Daily News (US). In total 4,439,440 articles were included in the analysis.

3.1 Translating text

We translate text using the traditional statistical phrase based machine translation [14] method with Moses. Statistical machine translation of text can be formulated as follows: Given a source sentence written in a foreign language f, Bayes rule is applied to reformulate the probability of translating f into a sentence written in a target language e.

$$e_{best} = argmax_e p(e|f) = argmax_e p(f|e) p_{LM}(e) \quad (1)$$

$p(f|e)$ is the probability of translating e to f and $p_{LM}(e)$ is the probability of producing a fluent sentence e. Sentences are broken in to phrases instead of words and phrases between the source and target language are aligned for training a translation model.

In our work translation models were trained with Moses for French, Spanish, German and Italian using the WMT (Workshop on Machine Translation) 2015 shared task training data. During the translation process Moses scores translation hypotheses using a linear model. Tuning refers to the process of finding the optimal weights for this linear model, where optimal weights are those which maximise translation performance on a small set of parallel sentences (Tuning set). We tuned our trained models with the tuning set from WMT using the MERT (Minimum error rate training [27]) algorithm. The output translation table from Moses contains all phrase pairs found in the parallel training corpus including a lot of noise. To reduce this noise the table is then pruned [28], resulting in faster loading of the model in memory. A language model was trained using all the available English corpora in WMT. We evaluated our trained models in the WMT test set using BLEU [12] score metric which is the most used metric based on n-gram precision computed between the machine generated translation and human generated translation. It ranges between 0-100%, and larger value identifies better translation. The idea behind BLEU is the closer a machine translation is to a professional human translation, the better it is. We obtain the following BLEU scores for each translation model: French (28.14%), Spanish (30.91%), German (26.11%) and Italian (30.69%). Translation models were deployed as Moses services, supporting multi-threading, so that all translation services run at a single point and several clients could request and collect translations at the same time. We translated news articles from all the outlets mentioned above including Euronews to English.

3.2 Measuring Sentiment

We used the LIWC sentiment word lists for positive and negative emotions (named as posemo, negemo) to measure sentiment about an entity in news articles. An entity in this context refers to named entities such as Persons, Organisations or Locations. Sentiment scores were computed in two weekly intervals with a one week overlapping time series window. For a given two weekly period, we obtain all news articles from the outlets mentioned above and search them for the given entity using Apache Lucene text search engine [15]. For each mention of the word we compute the number of positive and negative words surrounding the word with a text window of size 5 and total them for the period. Sentiment scores *psr* and *nsr* were computed. *psr* refers to positive sentiment ratio which is the number of positive words (p) divided by the total words in sentences containing the entity. The negative sentiment ratio (*nsr*) is computed similarly using the number of negative words (n).

We also measure the sentiment distance for an entity given by $(p-n)/(p+n)$. It shows how negative or positive was the news coverage about the entity in a given period of time.

3.3 Measuring Attention

The relative attention of an entity is computed by counting the number of entity mentions in a given period, dividing it by the total number of words from all articles in that period. This is again computed in two weekly intervals with a one week overlapping time series window.

4 Measuring Sentiment in Machine Translated Text

In this section we discuss and validate how machine translated text can be used to measure the sentiment across different languages. We compare the sentiment about two chosen entities 'Europe' and 'Russia', extracted from the English version with that extracted from the machine-English (translated) versions, reporting significant correlation both in positive sentiment and negative sentiment. For this purpose we use English, French, German, Spanish and Italian news articles from Euronews from 2015 to 2017. Euronews is a multilingual news media service publishing news content in several languages other than English. For each non-English article there is an equivalent English article published.

We translated Euronews news articles in French, German, Spanish and Italian to English using our trained models. Starting from the machine-English article, for each language we match its equivalent original English article by computing cosine similarities between the document vectors (term frequency vectors) across all relevant pairs obtained by date of article. We only filter the pairs that have a cosine similarity $(\theta) > 0.5$ for computing sentiment correlation between pairs for entities 'Europe' and 'Russia'. For each mention of the entity in the pair of articles, we compute the number of positive and negative words surrounding the entity with a text window size 5 and calculate scores *psr* and *nsr* as described in section 3.2. We compute the Pearson correlation coefficient between *psr* and *nsr* vectors for the English and machine-English pairs resulting in a positive and negative correlation score for the words in each language. Table 1 shows for each language, the total number of matching pairs found, then for each topic, it shows the total pairs of articles containing the topic, the total number

Table 1: Positive and Negative Sentiment correlation scores and p-values for null hypothesis between machine-English and original English article pairs for entities 'Europe' and 'Russia'

Language	Total pairs	Topic	Total pairs with topic	Total pairs with topic $\theta > 0.5$	Positive corr (p-val)	Negative corr (p-val)
French	14644	Europe	1113	219	0.25(0.03)	0.59(0.0)
		Russia	1190	118	0.13(0.03)	0.41(0.05)
German	15849	Europe	1886	541	0.24(0.0)	0.62(0.0)
		Russia	1428	346	0.31(0.02)	0.14(0.08)
Italian	16886	Europe	2056	458	0.50(0.0)	0.47(0.0)
		Russia	1523	255	0.53(0.03)	0.33(0.01)
Spanish	16821	Europe	2337	738	0.38(0.0)	0.60(0.0)
		Russia	1568	424	0.22(0.06)	0.51(0.0)

of pairs containing the topic with a cosine similarity $(\theta) > 0.5$ and positive and negative correlation scores for these pairs. Overall we observe positive and negative correlation in the region of 13%-53% and 15%-63%. To test the null hypothesis, in each language from the pairs with $(\theta) > 0.5$ we randomly assign machine-English articles to English articles, form pairs and compute the positive and negative correlation scores like before in 100 iterations. Our test statistic is the correlation score obtained and in each iteration we check if the correlation score is greater than the actual score obtained for each word in each language. Our p-value is $n/100$ where n is the number of times the score was greater than or equal to the actual score obtained. Table 1 shows the corresponding p-values next to the actual score. The p-values are very low ranging from 0.0 to 0.08. Therefore we conclude that the observed correlation is significant and sentiment can be measured across languages using machine translation.

5 Results and Discussion

In this section we demonstrate how sentiment and attention has changed in different periods of time over the past few years for entities Weinstein, Trump and Russia comparing the trends across US, UK, French, Spanish, German and Italian (FSGI) news outlets. The timelines were computed for each region by summing up the following counts across all outlets from that region in a two weekly period with a one week overlapping time series window. The counts are the number of positive (p), negative (n) words surrounding an entity, the total number of words in sentences containing the entity and the total number of words in all sentences. These counts were used to compute the positive sentiment ratio (psr), negative sentiment ratio (nsr), sentiment distance and attention towards an entity as discussed in section 3. We also measure the correlation between the sentiment distance series and attention series for the entities in US-UK, US-FSGI and UK-FSGI regions.

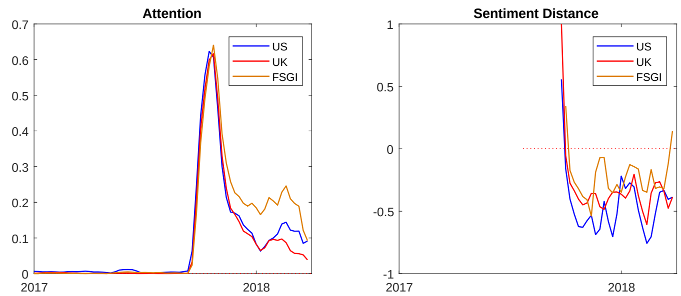


Fig. 1: The Attention and Sentiment Distance for the entity Weinstein in US, UK and FSGI news outlets starting from January 2017 until March 2018 analysing 716,610 articles. Period from January to September, 2017 is not shown in sentiment plot due to high error bars.

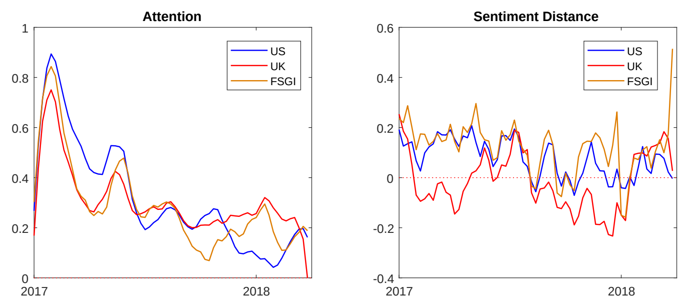


Fig. 2: The Attention and Sentiment Distance for the entity Trump in US, UK and FSGI news outlets starting from January 2017 until March 2018 analysing 716,610 articles.

5.1 Shifts in Sentiment and Attention

Harvey Weinstein Scandal The Harvey Weinstein scandal first came to light on the 5th of October, 2017 in a New York Times article[1], publishing a story detailing allegations of sexual harassment against him. Since then there have been a larger number of women coming forward to confess that they were harassed by him. We plot two quantities in Figure 1, showing the attention and sentiment distance for Weinstein in US, UK and FSGI new outlets starting from January 2017 until March 2018. The period from January to September, 2017 is not shown in sentiment plot due to high error bars. Error bars were calculated based on Wilson score confidence interval [17]. A high error bar indicates that the sentiment score for that period was not supported by enough positive and negative mentions about the entity.

[1] https://www.nytimes.com/2017/10/05/us/harvey-weinstein-harassment-allegations.html

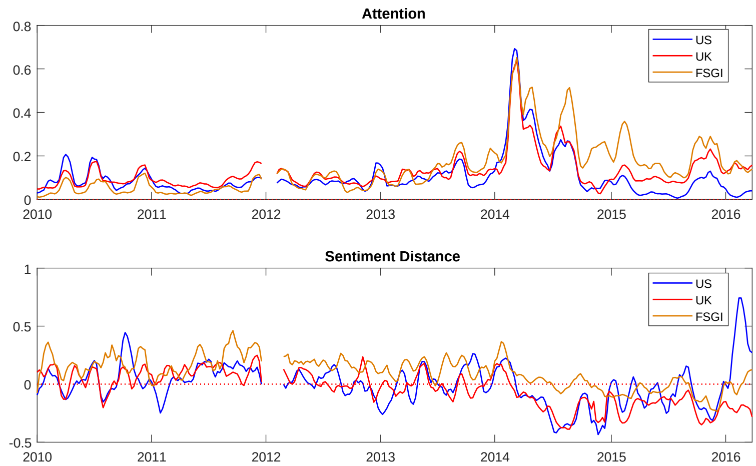
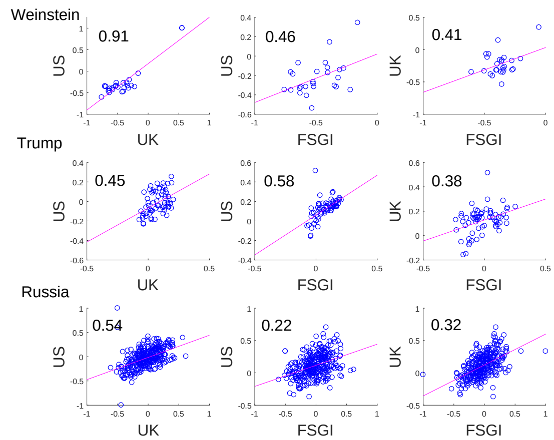


Fig. 3: The Attention and Sentiment Distance for the entity Russia in US, UK and FSGI news outlets starting from January 2010 until March 2016 analysing 3,722,830 articles. The period beginning in 2012 shows NaN values due to missing data.

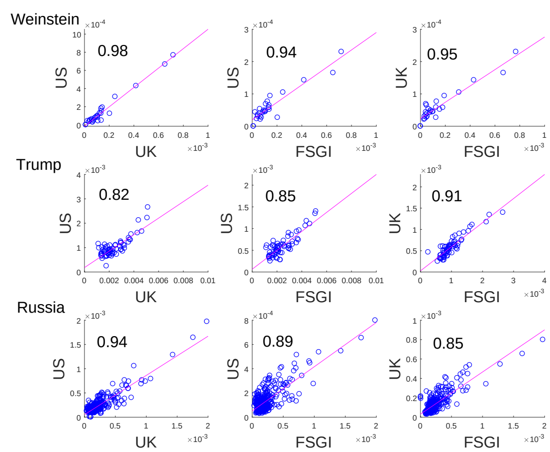
We see that the coverage starts from the beginning of October, 2017 with a massive peak in all three regions marking the outbreak of the scandal on 5th October. News media in all three regions show a steady increase in negative sentiment from the beginning of October and continues to be negative throughout the latter part of 2017 and 2018. The sentiment is more positive in FSGI and more negative in UK an US, although it is overall negative in all regions.

The correlation for attention is very high for Weinstein according to the correlation plot in Figure 4 in all three regions while Sentiment distance is much more correlated in US-UK than US-FSGI or UK-FSGI meaning the way the story was covered in the US and UK is very similar.

Donald Trump's Reactions to Riots Donald Trump's presence in the media became increasingly negative from August 2017 which continued until January 2018. Fig 2 shows the Attention and Sentiment Distance for the entity Trump in US, UK and FSGI new outlets starting from 2017 January until March 2018. We see that Trump was discussed more in US than UK or FSGI in the beginning of 2017. Sentiment clearly shows a negative shift happening in the news coverage of US, UK and FSGI during the first weeks of August. This was due to the reaction of Trump to the Charlottesville Riots on 12th August 2017 [30], where white supremacists clashed with counter-demonstrators during a rally. It was followed by a period of negative coverage in news for Trump when he repeatedly



(a) Sentiment distance



(b) Attention

Fig. 4: Correlation plots of Sentiment distance series and Attention series between, US-UK, US-FSGI and UK-FSGI for Weinstein (October 2017 - Mar 2018), Trump (Jan 2017 - Mar 2018) and Russia (Jan 2010 - Mar 2016) monitored every two weeks with a one week overlapping time series window.

criticized the NFL players who protested against the national anthem to bring attention to racial injustice in the United States [29].

The UK has the most negative coverage of Trump from the beginning of the riots. We see at periods it is overall negative, whereas in other EU countries and US it is more balanced. The correlation for media attention is high for all three regions according to Figure 4 while the Sentiment distance is also correlated across regions, more in US-FSGI.

Russia's Military Intervention in Ukraine The Russian military intervention in Ukrainian territory started in February 2014. The Crimean peninsula was annexed from Ukraine by the Russian Federation in February-March 2014 [31]. We see how this event has caused a sentiment shift for Russia in global media. Figure 3 shows the Attention and Sentiment Distance for the entity Russia in US, UK and FSGI new outlets starting from 2010 January until March 2016.

Coverage of Russia peaks in February 2014 Crimean crisis, attracting negative coverage in all regions and very negative in US and UK than FSGI. A rise in attention is observed again in July 2014 while an increasingly negative coverage is also observed at the same time during the shot down of Malaysia Airlines Flight 17 over an area of Ukraine. The attention seems to be higher in FSGI from this period. The Russian intervention in Crimea and Eastern Ukraine caused a lot of reactions and this period of negativity is well evident in the sentiment distance plot showing negative coverage about Russia in the US, UK and FSGI regions in the period between February 2014 and January 2016.

Correlation of media attention is very high for Russia across all three regions and Sentiment distance is more correlated in US-UK than others.

5.2 Conclusions

Monitoring the contents of the global media is an important task that requires automation. Challenges range from the correct way to measure and validate sentiment, to the problem of operating across languages. In this study we have demonstrated that it is possible to extract usable sentiment signals from machine-translated text, in a news setting, and we have presented a study of how media sentiment has changed across UK, US and some European counties, over a long time period. There are significant correlations across the three regions, but also interesting differences. Along the way we have also proven that sentiment can be measured across translation by creating a new aligned corpus of news that are paired using the translated Euronews articles from French, German, Spanish, Italian and their equivalent English article in Euronews where we assume that the sentiment of the coverage of a given entity is similar. We found that the positive and negative sentiment between these pairs are correlated showing that we can measure sentiment across languages through machine translation.

Acknowledgements. Saatviga Sudhahar and Nello Cristianini are supported by the ERC Advanced Grant "ThinkBig awarded to NC.

References

1. Tribou A. and Collins K. (2015) 'This Is How Fast America Changes Its Mind', Bloomberg, 26 June [Online]. Available at: https://www.bloomberg.com/graphics/2015-pace-of-social-change/

2. Silver N. (2015) 'Change Doesnt Usually Come This Fast', FiveThirtyEight, 26 June [Online]. Available at: https://fivethirtyeight.com/features/change-doesnt-usually-come-this-fast/
3. Lansdall-Welfare, T., Sudhahar, S., Veltri, G. A., & Cristianini, N. (2014, October). On the coverage of science in the media: A big data study on the impact of the Fukushima disaster. In Big Data (Big Data), 2014 IEEE International Conference on (pp. 60-66). IEEE.
4. Maynard, D., Gossen, G., Funk, A., & Fisichella, M. (2014). Should I care about your opinion? Detection of opinion interestingness and dynamics in social media. Future Internet, 6(3), 457-481.
5. Mutz, D., & Soss, J. (1997). Reading Public Opinion: The Influence of News Coverage on Perceptions of Public Sentiment. The Public Opinion Quarterly, 61(3), 431-451.
6. Su, L. Y. F., Cacciatore, M. A., Liang, X., Brossard, D., Scheufele, D. A. & Xenos, M. A. (2017). Analyzing public sentiments online: Combining human-and computer-based content analysis. Information, Communication & Society, 20(3), 406-427.
7. Young, L., & Soroka, S. (2012). Affective news: The automated coding of sentiment in political texts. Political Communication, 29(2), 205-231.
8. Nyman, R., Kapadia, S., Tuckett, D., Gregory, D., Ormerod, P., & Smith, R. (2018). News and narratives in financial systems: exploiting big data for systemic risk assessment. Bank of England Working Paper No. 704. Available at SSRN: https://ssrn.com/abstract=3135262 or http://dx.doi.org/10.2139/ssrn.3135262
9. McLaren, L., Boomgaarden, H., & Vliegenthart, R. (2017). News Coverage and Public Concern About Immigration in Britain. International Journal of Public Opinion Research, edw033.
10. Koehn, P., Hoang, H., Birch, A., Callison-Burch, C., Federico, M., Bertoldi, N., ... & Dyer, C. (2007, June). Moses: Open source toolkit for statistical machine translation. In Proceedings of the 45th annual meeting of the ACL on interactive poster and demonstration sessions (pp. 177-180). Association for Computational Linguistics.
11. Pennebaker, J. W., Francis, M. E., & Booth, R. J. (2001). Linguistic inquiry and word count: LIWC 2001. Mahway: Lawrence Erlbaum Associates, 71(2001), 2001.
12. Papineni, K., Roukos, S., Ward, T., & Zhu, W. J. (2002, July). BLEU: a method for automatic evaluation of machine translation. In Proceedings of the 40th annual meeting on association for computational linguistics (pp. 311-318). Association for Computational Linguistics.
13. Flaounas, I., Ali, O., Turchi, M., Snowsill, T., Nicart, F., De Bie, T., & Cristianini, N. (2011, June). NOAM: news outlets analysis and monitoring system. In Proceedings of the 2011 ACM SIGMOD international conference on Management of data (pp. 1275-1278). ACM.
14. Koehn, P., Och, F. J., & Marcu, D. (2003, May). Statistical phrase-based translation. In Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology-Volume 1 (pp. 48-54). Association for Computational Linguistics.
15. Jakarta, A. (2004). Apache Lucene-a high-performance, full-featured text search engine library. Apache Lucene.
16. O'Connor, B., Balasubramanyan, R., Routledge, B. R., & Smith, N. A. (2010). From tweets to polls: Linking text sentiment to public opinion time series. Icwsm, 11(122-129), 1-2.

17. Wilson, T., Hoffmann, P., Somasundaran, S., Kessler, J., Wiebe, J., Choi, Y., ... & Patwardhan, S. (2005, October). OpinionFinder: A system for subjectivity analysis. In Proceedings of hlt/emnlp on interactive demonstrations (pp. 34-35). Association for Computational Linguistics.
18. Gonzalez-Bailon, S., Banchs, R. E., & Kaltenbrunner, A. (2010). Emotional reactions and the pulse of public opinion: Measuring the impact of political events on the sentiment of online discussions. arXiv preprint arXiv:1009.4019.
19. Lin, Y. R., Margolin, D., Keegan, B., & Lazer, D. (2013, May). Voices of victory: A computational focus group framework for tracking opinion shift in real time. In Proceedings of the 22nd international conference on World Wide Web (pp. 737-748). ACM.
20. Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market. Journal of computational science, 2(1), 1-8.
21. Balahur, A., & Turchi, M. (2012, July). Multilingual sentiment analysis using machine translation?. In Proceedings of the 3rd workshop in computational approaches to subjectivity and sentiment analysis (pp. 52-60). Association for Computational Linguistics.
22. Balahur, A., & Turchi, M. (2014). Comparative experiments using supervised learning and machine translation for multilingual sentiment analysis. Computer Speech & Language, 28(1), 56-75.
23. Banea, C., Mihalcea, R., Wiebe, J., & Hassan, S. (2008, October). Multilingual subjectivity analysis using machine translation. In Proceedings of the Conference on Empirical Methods in Natural Language Processing (pp. 127-135). Association for Computational Linguistics.
24. Bautin, M., Vijayarenu, L., & Skiena, S. (2008, April). International Sentiment Analysis for News and Blogs. In ICWSM.
25. Bradley, M.M., & Lang, P.J. (1999). Affective norms for English words (ANEW): Stimuli, instruction manual and affective ratings. Technical report C-1, Gainesville, FL. The Center for Research in Psychophysiology, University of Florida.
26. Flaounas, I., Lansdall-Welfare, T., Antonakaki, P., & Cristianini, N. (2014). The anatomy of a modular system for media content analysis. arXiv preprint arXiv:1402.6208.
27. Och, F. J. (2003, July). Minimum error rate training in statistical machine translation. In Proceedings of the 41st Annual Meeting on Association for Computational Linguistics-Volume 1 (pp. 160-167). Association for Computational Linguistics.
28. H. Johnson, J. Martin, G. Foster & R. Kuhn. (2007) "Improving Translation Quality by Discarding Most of the Phrasetable"'. In Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL), pp. 967-975.
29. Malyon E., (2017) 'Donald Trump attacks 'disrespectful' NFL decision to allow players to continue protests against racial inequality', 18 October [Online]. Available at: https://www.independent.co.uk/sport/us-sport/national-football-league/donald-trump-nfl-players-protests-racial-inequality-kneel-anthem-colin-kaepernick-a8006806.html
30. Jacobs B. & Laughland O. (2017) 'Charlottesville: Trump reverts to blaming both sides including 'violent alt-left', 16 Aug [Online]. Available at: https://www.theguardian.com/us-news/2017/aug/15/donald-trump-press-conference-far-right-defends-charlottesville
31. Wikipedia contributors (2014) 'Annexation of Crimea by the Russian Federation', (Mar,2014) [Online]. Available at: https://en.wikipedia.org/wiki/Annexation_of_Crimea_by_the_Russian_Federation