Political News Propagation Headline Matching

A Deep Learning-Based Short Text Matching Approach

Zara Ahmad-Post
Management Information Systems
University of Arizona
Tucson, Arizona
zahmadpost@email.arizona.edu

Sagar Samtani
Operations and Decision Technologies
Indiana University
Bloomington, Indiana
ssamtani@iu.edu

Hongyi Zhu
Information Systems and Cyber Security
University of Texas San Antonio
San Antonio, Texas
hongyi.zhu@utsa.edu

Sue Brown
Management Information Systems
University of Arizona
Tucson, Arizona
suebrown@arizona.edu

ABSTRACT

Russian computational propaganda techniques have evolved to include the use of English-language news sites that serve as proxies for Russian propaganda. Information sharing and content propagation studies currently examine exact article copies but do not capture key semantics employed by proxy sites. We propose a short text matching deep learning (DL)-based model that seeks to capture the semantic similarity between article headlines in different media ecosystems. Our model captures named entities, word co-occurrences, and phrase sequences, and represents them in a propagandist context to capture both lexical and semantic similarities between article headlines. We benchmark our model against DL-based models from several architecture categories: Deep Neural Networks, Convolutional Neural Networks, Long-Short-Term-Memory Networks, and Attention-based architectures. Our proposed model consistently outperforms benchmark models by statistically significant margins across two datasets. Our results indicate that using several representations of text, combined with contextualization of content, can identify semantically similar article headlines with greater precision than some prevailing models.

CCS CONCEPTS

• Novelty in Information Retrieval • Social Aspects in Security and Privacy

KEYWORDS

Disinformation, deep learning, short text matching, artificial intelligence, cybersecurity, analytics

ACM Reference format:

Redacted for double blind review. 2021. Political News Propagation Headline Matching: A Deep Learning-Based Short Text Matching Approach. In *Proceedings of ACM KDD Workshop on Artificial Intelligence-enabled Cybersecurity. ACM, Singapore, 4 pages.* https://doi.org/10.1145/1234567890

1 Introduction

As the political disinformation and computational propaganda landscape continues to grow and evolve, so too have the tactics, techniques, and procedures (TTPs) used by disinformation and propagandist actors. Foreign nation-state actors' efforts to polarize U.S. politics and sow division amongst American citizens have been well documented [1, 2, 3, 4, 5]. A key adversary in the computational propaganda and political disinformation space is Russia, where Kremlin-backed efforts to influence the outcomes of U.S. democratic processes have played out to varying degrees of success. A report released by the U.S. Department of State in September, 2020 detailed a new tactic used by the Kremlin to disseminate messaging: the use of English-language proxy sites to spread propaganda [4]. These proxy sites are the "connective tissue" that disseminate propaganda from the Kremlin to U.S. audiences through news media ecosystems [6]. **Figure 1** illustrates how proxy sites affect propaganda dissemination.

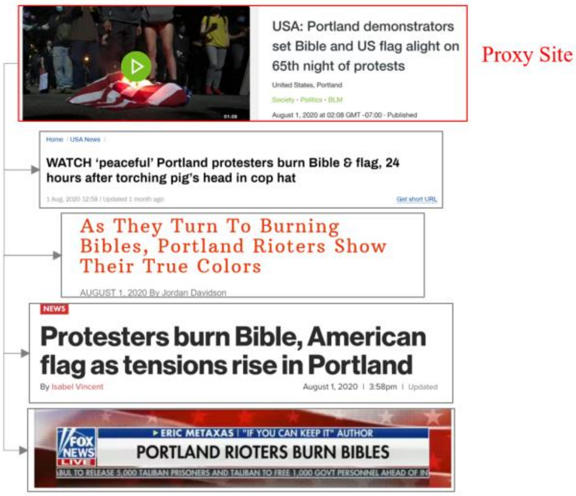

Figure 1: Semantic shifts in article headlines

The narrative illustrated in Figure 1 began when a Portland-based CBS affiliate published in a blog post in early August, 2020 that a protestor used a Bible and a flag to start a bonfire. The blog then states that other protestors quickly put the fire out. This story was picked up by a proxy site. As often seen in propaganda [7], the narrative undergoes subtle semantic shifts as the story propagates through news media ecosystems. For example, the words "protestors" and "rioters" are used interchangeably depending on the source; the presence of "American flag" or "US flag" is variable, and additional details are added in different headlines, (e.g., "hours after torching pig's head in cop hat," "tensions rise in Portland," "Rioters Show Their True Colors"). The fluctuation of semantics, combined with linguistic techniques frequently seen in propaganda [8], and noisy, variable text hinders conventional authorship attribution and text matching approaches. A suitable approach to capturing the semantic shifts in the headlines would need to include named entity recognition (e.g., to understand "Portland" as a city), word co-occurrences (e.g., "burned," "burning"), and phrase sequences (e.g., "Rioters burn Bible," "Protestors burn Bibles").

In this work, we develop a novel deep learning (DL)-based architecture for short text matching to explore how Russian propagandist narratives disseminate into mainstream or near-mainstream U.S. news ecosystems. Our proposed model captures named entities, word co-occurrences, and phrase sequences in article headlines, then represents these features in a propagandist context to better capture the semantic similarity between headlines from propagandist and non-propagandist sources. By capturing several representations of text and mapping the embeddings to a propagandist context, we can better match semantically similar, but lexically divergent, article headlines. As a result, the proposed model can match variable text in article headlines to map the flow of propaganda across news media ecosystems.

2 Related Work

2.1 Content Sharing in News Media Ecosystems

The domain of Russian propaganda has been well documented by intelligence communities and political science researchers [9, 10, 11]. Past studies provide deep historical and modern perspectives of Russian information operations, however they are largely qualitative analyses and are therefore not scalable.

Several recent studies have investigated content sharing in news media ecosystems. One study analyses content sharing across alternative and mainstream news media using a case study on the Syrian White Helmets (WH) [12]. The authors create content similarity scores of news articles sourced from Tweets mentioning WH using cosine similarity on term frequency-inverse document frequency (TF-IDF) word representations of the articles. Using a threshold of > .85 similarity, the authors capture near-exact matches of article content. Similarly, Horne et al. [13] examine content sharing practices across alternative, mainstream, and international news media ecosystems. They deploy similar methodologies to Starbird et al. [12], but do not limit their scope to case study-related content. Both studies utilize network analysis to create content sharing networks and perform further qualitative analyses to reveal content sharing ecosystems. Although both studies note the importance of Russian sources in these networks, neither study focuses specifically on Russian propaganda analysis. Furthermore, these studies use cosine similarity on TF-IDF representations of full articles to derive content similarity scores.

The studies focusing on exact content reproduction between and within news media ecosystems help identify newswire-like story sharing [12] but may not capture subtle semantic shifts between articles. One key adversarial advantage of Russian propaganda proxy sites is the creation of plausible deniability of connection to the Kremlin [4]. In this case, article reproductions are unlikely to be exact copies, but will still carry similar messaging as Russian state media sources. As these propagandist proxy sites are targeted to Western readers, it is important to design a method that can capture the properties of articles that appeal to these readers. For example, The American Press Institute notes that 60% of American readers only read headlines [14]. Headlines contain key messaging in both propagandist and quality journalism, and are excellent candidates for semantic matching between sources. Headlines are comprised of short text, the analysis of which is well suited to DL-based methods [15].

2.2 Deep Learning Short Text Matching Models

Prior literature that investigates content sharing in U.S. alternative and mainstream media utilize traditional text matching methods, such as computing cosine similarity on TF-IDF word vectors to derive a content similarity score. This approach captures exact article copies but not semantic shifts between article headlines. Moreover, propagandist text often includes exaggerated, loaded, and sometimes coded language to spread messaging [8]. Traditional text matching methods suffer in this context due to word variations and noise.

Short text DL-based approaches can address these issues. DL-based models facilitate semantic matching by projecting text into representations in a latent sematic space with multiple layers of non-linear activation functions, error correction, and backpropagation [16]. Four categories of DL-based short text matching methods exist:

- Deep Neural Network (DNN)-based models such as Deep Relevance Matching Model (DRMM) [17] stack dense layers to learn feature representations of data for matching [15].
- Convolutional Neural Network (CNN)-based models such as Kernel-based Neural Ranking Model (KNRM) [18] capture local text sequential dependencies with convolution [15].
- Long-Short Term Memory (LSTM)-based models such as MV-LSTM [19] utilize recurrent network architecture to capture sequential and temporal data patterns in text [15].
- Attention-based models such as Attention-based Neural Matching Model (aNMM) [20] use shared-value weighting schemes in deep neural networks to help determine the most relevant words in a query or document [20].

Despite the widespread usage of these techniques for many short text matching contexts, they were not designed to operate on propagandist headlines. As a result, these conventional DL-based short text matching algorithms cannot capture information essential to matching propagandist article headlines. For example, named entities need to be correctly identified to ensure semantic similarity is captured across texts. Word co-occurrences and phrase sequences need to be captured. Finally, propagandist language needs to be contextualized to properly identify semantic matches in the corpus.

2.3 Research Gaps and Questions

We identified several research gaps from our literature review. First, past work that focus specifically on Russian propaganda are largely qualitative and are therefore not scalable. Second, past studies focusing on content sharing practices in news media ecosystems do not capture subtle semantic shifts in messaging across sources. Finally, although DL-based approaches are well suited to this task, to best of our knowledge, no models current exist that incorporate the necessary components of article headlines to facilitate propagandist article headline matching. Given these research gaps, we propose the following research questions for study:

- How can we use article headline data to map the dissemination of Russian propagandist news stories and narratives to American news media ecosystems?
- How can we capture the named entities, word co-occurrences, and phrase similarities of article headlines in a DL-based short text matching model?

3 Methods

3.1 Data

Using RSS feeds, we collected the 26,234 articles (headlines, body text, publication dates and times, and authorship information) since 11/2/2020 from 24 sites. Sites in our collection include Russian state media, Russian propaganda proxy sites, and a wide range of U.S. mainstream and alternative media. Russian state media sites targeted for collection include RT, Sputnik News, and ITAR-TASS. Russian propaganda proxy sites include Strategic Culture Foundation, Global Research, Katehon, Geoplotica.ru, Ruptly, and NewsFront. These sites were identified by the U.S. Department of State as Russian propaganda proxy sites [4]. U.S. mainstream and alternative media sites collected include the New York Post, Fox News, The Washington Post, USA Today, the Star Tribune, the Chicago Tribune, The LA Times, The New York Times, HuffPost, 21st Century Wire, Blacklisted News, One America News Network, Breitbart, Info Wars, and Newsmax. Also, each site collected is standalone, full-featured, and Western readers-focused, which publishes and regularly updates their RSS feeds and contains reporting that is predominantly textual.

3.2 Gold-Standard Datasets

Short text matching models require labeled pairs of relevant and irrelevant texts to train the model [16]. Using prta [21], we locate all potential Russian propagandist headlines in our dataset. For each Russian propagandist headline, five candidate headline matches were located by computing cosine similarity on Word2Vec vector representations of article headlines, then manually verified. If one of the five headlines carried the same meaning as the propagandist headline, it was labeled as a match. We then randomly selected four headlines from the dataset for each propagandist headline, manually verifying that they were semantically dissimilar to the query headlines. Utilizing this method, we created two gold-standard datasets: one with headlines related to the COVID-19 pandemic and vaccinations and the other with headlines related to international affairs and topics. **Table 1** summarizes our gold-standard datasets.

Topic	COVID-19	International Affairs
Total # Pairs	500	325
Training (80%)	400	260
Validation (10%)	50	33
Testing (10%)	50	32

Table 1: Summary of Gold Standard Datasets

3.2 Proposed Model

We propose a novel DL-based architecture that captures named entities, word co-occurrences, and phrase similarities between propagandist and non-propagandist texts. We present our proposed model in **Figure 2.**

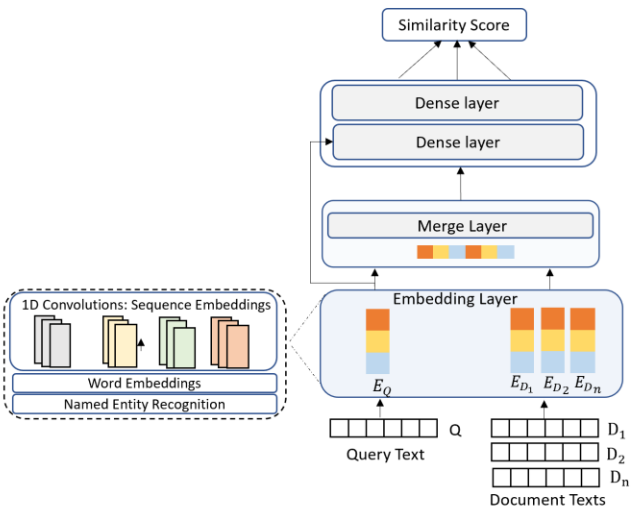

Figure 2: Proposed Model Architecture

The proposed model operates as follows:

1. The inputs to the model are pairs of article headlines: a query headline Q, and n document headlines $D_1, D_2, \dots, D_n$.
2. Each query and document headline is first parsed to produce named entity tokens $NET(Q)$ and $NET(D_i)$. 18 different entity types are recognized by spaCy [22], a common Python natural language processing library. Therefore, named entities tokens are represented as binary values in a $W \times 18$ matrix, where W is the headline length.
3. Consistent with best practices, we concatenate Word2Vec embeddings [23] to represent each query and document. The output $WE(Q)$ and $WE(D_i)$ are $W \times 300$ matrices.
4. The resulting matrix is then passed to a 1-dimensional (1D) convolutional layer with window size=3. This layer captures word trigrams to represent phrase sequences. The output of the 1D convolutional layer is a $W \times 300$ matrix for each query and document text.

$$SEQ(Q) = Conv1D_Q(Q),\ SEQ(D_i) = Conv1D_D(D_i).$$

5. The named entity recognition values, word embeddings, and sequence embeddings are concatenated to form the full embeddings for each query as E_Q and document text as E_{D_i}.
6. The resulting document embedding is then fed to a merging layer. This layer concatenates the query and document embeddings and passes them to a dense layer, which serves to represent the document embedding in terms of the query embedding. This mechanism transforms the document embedding into a representation that captures elements of the propagandist text.

$$D_i^M = Dense_M([E_Q; E_{D_i}])$$

7. Query and document embeddings are fed to two dense layers, which perform dimensionality reduction.

$$Q' = Dense_2(Dense_1(E^Q)), D_i' = Dense_2\left(Dense_1(D_i^M)\right)$$

8. A final layer computes a cosine similarity score between the query and the document.

$$Output = \left(sim(Q^i, D_1'), sim(Q^i, D_2'), \dots, sim(Q^i, D_n')\right)$$

The design novelty of our proposed algorithm resides in the representation of the text, which captures named entities, word co-occurrences, and phrase sequences, as well as the contextualization of document text in terms of propagandist query text. The benefits of this approach are twofold. First, we reduce reliance on a single embedding or text representation, which may falter in noisy short text. Second, by leveraging a merge layer, we avoid feature engineering (manual, time-intensive, error-prone) that represent propagandist techniques (e.g., appeals to authority, loaded language, etc.)

4 Results and Discussion

We benchmark our proposed model against DL-based models in each major algorithm architecture category (DNN, CNN, LSTM, and attention-based models). We evaluate each model's performance on our gold-standard datasets using ranking metrics common in information retrieval tasks. Consistent with best practice in short text matching literature, we measure the Mean Average Precision (MAP), Mean Reciprocal Rank (MRR), and Normalized Discounted Cumulative Gain (NDCG) at 1, 3, and 5 [24]. We present our results in **Table 2**.

Data Set	**COVID-19**				
Metric	**MAP**	**MRR**	**NDCG@1**	**NDCG@3**	**NDCG@5**
DRMM	0.448***	0.448***	0.211***	0.395***	0.566***
MVLSTM	0.662	0.675	0.466	0.699	0.744
KNRM	0.548***	0.556***	0.300**	0.570***	0.651***
aNMM	0.507***	0.511***	0.288***	0.474***	0.615***
Proposed Model	**0.698**	**0.698**	**0.544**	**0.710**	**0.773**
Data Set	**International Affairs**				
Metric	**MAP**	**MRR**	**NDCG@1**	**NDCG@3**	**NDCG@5**
DRMM	0.423***	0.423***	0.200***	0.413***	0.529***
MVLSTM	0.458**	0.458**	0.285*	0.423**	0.555***
KNRM	0.430***	0.430***	0.214***	0.413***	0.534***
aNMM	0.426***	0.426***	0.214**	0.411***	0.532***
Proposed Model	**0.649**	**0.649**	**0.457**	**0.654**	**0.736**

*p-value<0.05, **p-value<0.01, ***p-value<.001

Table 2: Experiment Results

Our model outperforms the best-competing benchmark model by .1714% in NDCG@1 in the international affairs dataset. Our proposed model outperforms the best-competing benchmark model by .0777% in the COVID-19 dataset, although it does not do so by statistically significant margins.

We present two query-document pairs that were returned correctly by our proposed model but were missed by the best-competing benchmark model (MVLSTM). In the first example, the query headline, “Hurricane Iota now lashing Nicaragua as a Category 2 storm, but it’s still a record-setter” is paired by our model with “Hurricane Eta slams into Nicaragua as a Category 4 storm, causing damage and rivers to overflow.” MV-LSTM returned the headline, “Monday Night Football Ratings Crash by 30%, ESPN's Least Watched Week 9 Game Since 2004.” These results indicate that our model is able to capture semantic similarity between article headlines, even if the explicit meaning is slightly different. Additionally, this example demonstrates that named entities (Nicaragua), word co-occurrences, (Category, storm), and phrase sequences (Hurricane…Nicaragua as a Category) are being correctly identified.

In the second example, the headline, “Secret Service Struck Again by Coronavirus Outbreak,” is correctly paired by our proposed model with the article headline “Dozens of Secret Service officers sidelined by coronavirus outbreak.” The MV-LSTM model returned the document, “Politics live updates: Congress to pick up COVID relief talks as lawmakers eye stimulus checks.” Although topically relevant, the headline returned by MV-LSTM is not a semantic match for the query headline. Our proposed model captures named entities (Secret Service), word co-occurrences (coronavirus outbreak), and phrase sequences (Secret Service… by coronavirus outbreak).

Overall, these results suggest that capturing named entities, word co-occurrences, and phrase sequences in propagandist text can identify semantic similarity between article headlines. Our approach enables better semantic matching of article headlines with variable lexical similarity. As a result, the proposed method could be an important tool to help trace the origins of propagandist narratives and disinformation.

6 Conclusions and Future Directions

Tracing the flow of propaganda through news media ecosystems is critical to understanding the sources of narratives. Critical to this task is the ability to match article headlines with subtle semantic variability. We have demonstrated that a model that captures named entities, word co-occurrences, and phrase co-occurrences, and represents documents in the context of queries is well suited to the task of matching semantically similar article headlines. Our proposed model outperforms benchmark DL-based models that do not incorporate these techniques.

Several promising directions for future work can incorporate network analytics to better understand headline propagation. Additionally, the impact of propagation of propaganda cannot be fully understood without including analysis of the reach of the content. However, this may be best understood through social network analysis of mainstream and fringe social networks.

REFERENCES

[1] Adam Badawy, Aseel Addawood, Kristina Lerman, *et al.* 2019. Characterizing the 2016 Russian IRA influence campaign. *Social Network Analysis Mining* 9, 31. DOI: https://doi.org/10.1007/s13278-019-0578-6.

[2] Gabriele Cosentino. 2020. Polarize and conquer: Russian influence operations in the United States. In *Social Media and the Post-Truth World Order*, 33-57. Palgrave Pivot, Cham.

[3] Elizabeth Anne Bodine-Baron, Todd C. Helmus, Andrew Radin, and Elina Treyger. 2018. Countering Russian social media influence. Rand Corporation, Santa Monica.

[4] Philip N. Howard, Samuel Woolley & Ryan Calo. 2018. Algorithms, bots, and political communication in the US 2016 election: The challenge of automated political communication for election law and administration. *Journal of Information Technology & Politics*, 15:2, 81-93, DOI: 10.1080/19331681.2018.1448735

[5] Department of State. 2020. GEC Special Report: Pillars of Russian Disinformation and Propaganda Ecosystem. Retrieved from https://www.state.gov/wp-content/uploads/2020/08/Pillars-of-Russia%E2%80%99s-Disinformation-and-Propaganda-Ecosystem_08-04-20.pdf

[6] Samantha Bradshaw, Phillip N. Howard. 2019. The Global Disinformation Order: 2019 Global Inventory of Organized Social Media Manipulation. Oxford Internet Institute, University of Oxford.

[7] Matthew Rosenberg, and Julian Barnes. 2020. A Bible Burning, a Russian News Agency, and a Story Too Good to Check Out. (August 11, 2020). *The New York Times*. Retrieved from https://www.nytimes.com/2020/08/11/us/politics/russia-disinformation-election-meddling.html.

[8] Clyde R Miller. 1939. How to detect and analyze propaganda. The Town Hall, Inc, New York City.

[9] Geir Hagen Karlsen. 2019. Divide and rule: Ten Lessons about Russian Political Influence Activities in Europe. *Palgrave Commun* **5,** 19. https://doi.org/10.1057/s41599-019-0227-8

[10] Marcel H. van Herpen. 2016. Putin's propaganda machine : soft power and Russian foreign policy. Rowman and Littlefield, Lanham, MD

[11] Peter Pomerantsev. 2015. “Authoritarianism Goes Global (II): The Kremlin’s Information War.” *Journal of Democracy* 26, 4, 40–50.

[12] Kate Starbird, Ahmer Arif, Tom Wilson, Katherine Van Koevering, Katya Yefimova, and Daniel Scarnecchia. 2018. Ecosystem or Echo-System? Exploring Content Sharing across Alternative Media Domains. Presented at *Eleventh International AAAI Conference on Web and Social Media* (ICWSM 2018), Stanford, CA.

[13] Benjamin D. Horne, Jeppe Nørregaard, and Sibel Adalı. 2019. Different Spirals of Sameness: A Study of Content Sharing in Mainstream and Alternative Media. *Proceedings of the International AAAI Conference on Web and Social Media*, *13*(01), 257-266. Retrieved from https://aaai.org/ojs/index.php/ICWSM/article/view/3227

[14] American Press Institute, Survey Research. 2019. How Americans get their news. (2019, June 11). Retrieved December 01, 2020, from https://www.americanpressinstitute.org/publications/reports/survey-research/how-americans-get-news/

[15] Bhaskar Mitra. and Nick Craswell. 2018. An introduction to neural information retrieval. Foundations and Trends in Information Retrieval. Now Publishers Inc.

[16] Po-Sen Huang, Xiaodong He, Jianfeng Gao, Alex Acero, Larry Heck. 2013. Learning Deep Structure Semantic Models for Web Search using Clickthrough Data. In *International Conference on Information and Knowledge Management*, (Oct 27- Nov 1, 2013). San Francisco, CA.

[17] Jiafeng Guo, Yixing Fan, Qingyao Ai, and W. Bruce Croft. 2016. A deep relevance matching model for ad-hoc retrieval. In *Proceedings of the 25th ACM international on conference on information and knowledge management*, 55-64.

[18] Chenyan Xiong, Zhuyun Dai, Jamie Callan, Zhiyuan Liu, and Russel Power. 2017. End-to-end neural ad-hoc ranking with kernel pooling.

In *Proceedings of the 40th International ACM SIGIR conference on research and development in information retrieval*, 55-64.
[19] Shengxian Wan, Yanyan Lan, Jiafeng Guo, Jun Xu, Liang Pang, and Xueqi Cheng. 2016. A deep architecture for semantic matching with multiple positional sentence representations. In *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 30, no. 1.
[20] Liu Yang, Qingyao Ai, Jiafeng Guo, and W. Bruce Croft. 2016. "aNMM: Ranking short answer texts with attention-based neural matching model." In *Proceedings of the 25th ACM international on conference on information and knowledge management*, 287-296.
[21] Giovanni Da San Martino, Shaden Shaar, Yifan Zhang, Seunghak Yu, Alberto Barrón-Cedeno, and Preslav Nakov. 2020. Prta: A system to support the analysis of propaganda techniques in the news. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, 287-293.
[22] Montani Honnibal, and Ines Montani. 2017. spaCy 2: Natural language understanding with Bloom embeddings, convolutional neural networks and incremental parsing.
[23] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. Distributed representations of words and phrases and their compositionality. arXiv:1310.4546.
[24] Bhaskar Mitra, Fernando Diaz, and Nick Craswell. 2017. "Learning to Match Using Local and Distributed Representations of Text for Web Search," in *Proceedings of the 26th International Conference on World Wide Web (WWW'17)*, 1291-1299.