

# Towards Bias Inducing Word Detection by Linguistic Cue Analysis in News Articles

Sora LIM<sup>†</sup>, Adam JATOWT<sup>†</sup>, and Masatoshi YOSHIKAWA<sup>†</sup>

<sup>†</sup> Graduate School of Informatics, Kyoto University  
Yoshida-honmach, Sakyo-ku, Kyoto, 606–8501, Japan

E-mail: <sup>†</sup>infosora@db.soc.i.kyoto-u.ac.jp, <sup>††</sup>adam@dl.kuis.kyoto-u.ac.jp, <sup>†††</sup>yoshikawa@i.kyoto-u.ac.jp

**Abstract** Biased news still exists even though the balance, fairness and accuracy are important qualities in news reporting. The bias in news causes political and social bipolarization and what is worse, it provides fodder for fake news which is one of the serious social problems. To tackle the problem, therefore, we identify the words which induce bias for a given news article by analyzing linguistic cues, such as factive verbs, implicative words, hedges, subjectivity words, etc., used in it. Our strategy to detect bias word effectively is to compare the word usage with of other news articles about an identical news event. In this process, we employ the concept of inverse document frequency (IDF), which is a well-known statistic in the information retrieval, as one of the features to represent the word in the news article. We also present a news bias detecting method in a hierarchical manner. We anticipate that our approach can be applied to other domains such as Wikipedia as well as newspapers where the use of unbiased words is advised.

**Key words** News Bias, Bias Detection, Linguistic Feature Analysis, IDF

## 1. Introduction

News has a large influence on viewpoint and attitude towards social issues including politics and economy. Therefore the fairness, accuracy and balance are regarded as key requirements in news reporting. However unfair and biased news still surround us. Furthermore, so-called “Filter Bubble” effect boosts news readers to consume only their favorite news media consistently [14]. It aggravates the news bias problem.

News bias typically arises from the word choice for describing the news event. For example, Table 1 shows two different news articles about “returns of US surveillance planes from Iraqi airspace.” The news article on the left side uses “complain”, on the other hand the news article on the right side uses “threaten” about Iraq’s action which caused the two plains returned. It assigns more negative angle against Iraq’s action by using the latter representation.

Two articles in Table 2 describe two Democrats’ actions for the rescindment decision of president Trump on Deferred Action for Childhood Arrivals (DACA) program. In this figure, two different words are used as modifiers of the word “immigrants” who were the beneficiaries of the DACA program, i.e., undocumented and illegal. The use of the word “illegal” degrades immigrants by loading more negativity than the use of “undocumented”. By using these words the articles imply their stance on the rescindment of DACA by Donald Trump.

Even though news articles deals with identical event, the articles deliver the event with the different nuances by using different word

U.N.Withdraws U-2 Planes	Iraq forces suspension of U.S. surveillance flights
WASHINGTON (AP)-U.N. arms inspectors said Tuesday they had withdrawn two U-2 reconnaissance planes over Iraq for safety reasons after Baghdad <b>complained</b> both aircraft were in the air simultaneously.	UNITED NATIONS (Reuters) - Iraqi fighter jets <b>threatened</b> two American U-2 surveillance planes, forcing them to return and abort their mission and return to base, senior U.S. officials said Tuesday.

Table 1 News articles about “returns of US surveillance planes from Iraqi airspace.” (left: AP and right: Reuters)

Pelosi and Schumer Say They Have Deal With Trump to Replace DACA	Schumer, Pelosi announce deal with Trump on Dreamers
WASHINGTON - Democratic leaders on Wednesday night declared that they had a deal with President Trump to quickly extend protections for young <b>undocumented immigrants</b> and to finalize a border security package that does not include the president’s proposed wall.	Democratic leaders emerged from a dinner meeting Wednesday night with President Trump to say they had worked out a deal to grant permanent protections to young <b>illegal immigrants</b> - without having to accept funding for the president’s proposed border wall.

Table 2 News articles about two Democrats’ actions for the rescindment decision of the president Trump on Deferred Action for Childhood Arrivals (DACA)

choices so that it causes that readers of particular newspaper have biased point of view. In this paper, the words which cause bias on the event are called as bias words in news articles. Our main goal is

to identify the bias words in the news articles automatically.

In news articles, however, to identify those bias words is challenging. The bias words make a fine distinction of nuance rather than an opposite meaning directly because in principle, news aims for factuality and impartiality to carry the banner for journalism [21].

Previous works which have tackled bias detection problem mainly regards news articles or the words independently without comparing news event context to investigate the word usage. They mainly focus on grammatical role and predefined dictionary-based features to represent words in news articles. But there are still limitations. The previous approach needs to extend the size of dictionary when precision is more required. Thus, even if every news article uses a word  $w$  for describing an object  $A$ , their approach regards the word  $w$  as a bias word. However, the word  $w$  might not be a bias word if it is used across whole articles. It commonly exists, such as in “encouraging good and punishing evil” case, including hurricane damage. In this paper, we employ relativity as a feature to represent a word used in news to figure out its biasing potential from the whole dataset point of view.

In this paper, we present a news bias detecting method given a news article based on a hierarchical structure which consists of top down approach and bottom up approach. We also suggest a news bias contrasting system which shows the most contrasting news content describing the identical news event but reveal most different point of view from the given article. It helps news readers to aware of biased contents injurious to balanced cognition of news event.

## 2. Related Works

Many researches made efforts to resolve news bias problem. However most of them have focused on the news diversification according to the content similarity. Park et al. [15] developed a news diversification system, named NewsCube, to mitigate bias problem by provide diverse information to users. They utilized the structure of news text to capture the differences of news aspects represented by several keywords.

In [16], comments of users on news articles are analyzed to identify news bias in terms of its political stance.

Hambourg et al. [5] presented a matrix-based news analysis to display various perspectives for the same news topic in a two-dimensional matrix.

Ogawa et al. [13] focused on describing way of the relationship between main participants in news articles to detect news bias. To catch describing way of the relationship, they expanded sentiment words in SentiWordNet [3] which is lacking in relationship distinction.

An et al. [2] revealed skewness of news outlets by analyzing their news contents spread throughout the tweeter.

H. M. Alonso et al. [1] tackled omissions between news statements which are similar but not identical. The omission occupies one category in news bias. They also revealed that features com-

prised of simple linguistic cues increased the omission classification results on a manually annotated corpus.

Some works focused on linguistic analysis for bias detection on text data. Recasens et al. [18] targeted bias words in Wikipedia pages. They utilized linguistic resources as features to detect bias words from the revised sentence history in Wikipedia.

Baumer, E. et al. [4] utilized the Recasens’ linguistic features to identify framing language in political news as well as some features from theoretical literature on framing. Their detecting tasks using the features achieved comparable performance with that of humans in terms of accuracy, F1, and recall, but weakness in precision. This paper also has a base on the use of linguistic resources introduced in the previous works to identify where news bias appears in a news article.

To enable users to compare multiple contents at one time, Nadamoto and Tanaka proposed a new type of web browser called Comparative Web Browser (CWB) [12]. Their system automatically synchronizes the passages of the relevant web pages which are similar each other so that the users compare the similar information easily. B-CWB [11] extended the CWB for news articles written in different languages. It analyzed two similar news pages that report the same event and discovered differences between them based on the content synchronization. Ma and Yoshikawa proposed a news comparison system called TVBanc [10]. They analyzed bias and diversity in news content by comparing topics and viewpoints. Their approach looks similar with ours but we study deep into news bias. We use a hierarchical way that narrows down bias searching space for the news document, not only for the document in itself.

## 3. Bias Word Representation

To detect news bias words computationally, we represent each word by following features. Here, two types of features are used. - 1) baseline feature from previous works, and 2) relative feature which is devised for the relative strength of bias words.

### 3.1 Baseline features

Baseline feature includes grammatical role of a word based on part of speech, whether a word is a named entity or not, document structure where a word occurs in a news document, linguistic cues such as subjectivity/objectivity [19], sentiment [9], verb characteristics [18] [6] [7] [8], as well as the word itself. Table 3 shows the baseline feature in detail. All these features in baseline feature set are document-dependent, which means each feature in this set only cares about word occurrences in a given document according to predefined dictionary.

### 3.2 Relative Features

To detect subtle bias in news article, we present to use relativity as a feature. This feature captures how much a word in a document has biasing potential by comparing its usage through the corpus. The basic idea of relative feature is to put word usage through the corpus together, and then, compare the frequency of the usage. For this

ID	Feature	Description
1	Word	Word $w$ itself
2	Stem	stem of $w$
3	POS	POS of $w$
4	Position	Position of $w$ in the sentence
5	Hedge	$w$ is in Hyland, 2005
6	Factive verb	$w$ is in Hooper, 1975
7	Assertive verb	$w$ is in Hooper, 1975
8	Implicative verb	$w$ is in Karttunen, 1971
9	Report verb	$w$ is in Recasens et al., 2013
10	Entailment	$w$ is in Berant et al., 2012
11	Strong subjective	$w$ is in Riloff and Wiebe, 2003
12	Weak subjective	$w$ is in Riloff and Wiebe, 2003
13	Positive word	$w$ is in Liu et al., 2005
14	Negative word	$w$ is in Liu et al., 2005
15	Bias Lexicon	$w$ is in Recasens et al., 2013
16	Named Entity	$w$ is Named Entity

Table 3 Baseline Features

process, we apply the concept of the inverse document frequency (IDF), which is a well-known statistic in information retrieval. Basically, IDF measures whether a term is common or rare across all documents in the corpus as follows.

$$idf(w) = \frac{|D|}{df(d, w)} \quad (1)$$

where  $|D|$  is the total number of documents in the corpus, and  $df$  is the number of documents where the word  $w$  appears. This feature has flexibility in terms that it is unrestricted by predefined dictionary.

We can expand the IDF concept to the other features described previously. For example, to check bias from assertiveness side for a specific news event by using the assertive word dictionary, we can set the relative assertiveness as shown in equation (2).

$$rf = \frac{|D|}{df(\text{assertive words}, d)} \quad (2)$$

By applying word2vec approach, we take the usage of semantically similar words into account beyond the dictionary existence as shown in (3).

$$rf = \frac{|D|}{df(w' \text{ that is similar word of } w, d)} \quad (3)$$

It is also defined by means of sentiment words (or subjectivity words), named entities in (4) and (5) as below.

$$rf = \frac{|D|}{df(w' \text{ having similar polarity of } w, d)} \quad (4)$$

$$rf = \frac{|D|}{df(w \text{ that is a named entity}, d)} \quad (5)$$

## 4. News Bias Detection

In this section, we present the process to detect bias word using the previously described features. We deal with news articles in a

hierarchical manner which consists of three levels, i.e., document, paragraph and word level, to analyze the existence of bias in it. Here are two ways to tackle bias detection on the hierarchical structure ‘Top down’ approach and ‘Bottom up’ approach. The main difference between them is based on how the approach looks at the bias in news articles. The top down approach is based on ‘*News bias is detected by distinction of the similarities.*’ And the bottom up approach is based on ‘*News bias is the result of aggregation of the biasing units.*’ In this section, we describe those two approaches in detail.

### 4.1 Top down approach

The main idea in the top down approach is that we narrow down the focusing area of the news document to detect hidden bias inducing words step by step. Each step filters out similarity and only considers the difference between the articles.

The document level detects which document shows different content compared to the whole document set. For a given news article  $d$ , it constructs the set of the news articles  $D$  which describes the same news event with  $d$ . Each news article is represented as its keywords. Then, the distribution of news set is analyzed in terms of the similarity distance of main aspects of the articles. Next, for the target news  $d$ , we check whether the news  $d$  is located within a dense region of the cluster or not under a threshold  $\delta$ . We call the dense region majority. If  $d$  belongs to the majority, it is not a biased article. Otherwise, it is regarded as a bias candidate document. From the output of the document level, therefore, we can check the bias induced by the differences of main aspects of news articles even the articles are dealing with the same event.

In the paragraph level, we only consider the majority and filter out the others to check the news bias in detail. From the majority, we investigate which paragraph contains similar information on surface, but different underneath with other paragraphs from different articles in the majority. Here the difference could be the nuance implying news outlet’s likes and dislikes for the main agent of the event or the event itself.

So in this step, the output is paragraph pairs indicating the same specific information from different news articles. The information from them looks same on the surface, but each of them causes user understanding differently for the event.

The word level highlights which word in the paragraph makes the difference, i.e., bias inducing words. From the paragraph pairs, the semantic differences are analyzed. It yields bias inducing words in the three different degrees, such as high, middle and low according to their biasing scores. In this step the proposed relative feature plays an important role to calculate the biasing scores.

### 4.2 Bottom up approach

The bottom up approach starts with detecting bias inducing words. It assumes that the more bias inducing words appear in a news article the more biased the article becomes. First, given the target news article  $d$ , it detects the most likely words to cause dif-

## Extracted Data

idf >= 50% | idf >= 70% | idf >= 90%

Title Federal judge blocks Trump ' s third travel ban

Text A federal judge on Tuesday largely blocked the Trump administration from implementing the latest version of the president ' s controversial travel ban , setting up yet another legal showdown on the extreme of the executive branch ' s powers when it comes to setting immigration policy

The decision from US . District Judge Derrick K . Watson in Hawaii is sure to be appealed , but for now , it means that the administration cannot restrict the entry of travelers from six of the eight countries that officials said were unable or unwilling to provide information that the United States wanted to vet the countries ' citizens

The latest ban was set to go fully into effect in the early hours of Wednesday , barring various types of travelers from Syria , Libya , Iran , Yemen , Chad , Somalia , North Korea and Venezuela . Watson ' s order stops it , at least temporarily , with respect to all the countries except North Korea and Venezuela

In a 40page decision granting the state of Hawaii ' s request for a temporary restraining order nationwide , Watson wrote that the latest ban " suffers from precisely the same maladies as its predecessor "

Watson also wrote that the executive order " plainly discriminates based on nationality " in a way that is opposed to federal law and " the founding principles of this Nation "

The White House said in a statement that Watson ' s " dangerously flawed " order " undercuts the President ' s efforts to keep the American people safe and enforce minimum security standards for entry into the United States "

(a) Washington Post

## Extracted Data

idf >= 50% | idf >= 70% | idf >= 90%

Title Federal Court Blocks Donald Trump ' s Latest Travel Ban

Text The International Refugee Assistance Project , the Iranian Alliances Across Borders and various individual plaintiffs , including the state of Hawaii , filed lawsuits in Hawaii and Maryland after the travel ban which would indefinitely ban citizens from Chad , Iran , Libya , North Korea , Somalia , Syria , Venezuela and Yemen from entering the country was announced on Sept . They saw this ban as yet another thinly veiled attempt to target Muslims six out of the eight countries on the list are Muslimmajority

" EO3 suffers from precisely the same maladies as its predecessor it lacks sufficient findings that the entry of more than million nationals from six specified countries would be " detrimental to the interests of the United States , " " the court order said . The ban also " plainly discriminates based on nationality in the manner that the Ninth Circuit has found antithetical to both Section 192 and the founding principles of this Nation "

Nonvisa holders from most of the countries listed have already been banned for the last several months , after the Supreme Court ruled to allow parts of a previous travel ban to go into effect in June The latest initiative removed Sudan from the list , and added Chad , North Korea and Venezuela . It was planned to go into effect Wednesday

" This is the third time we feel compelled to ask the courts to block the President ' s Muslim ban , " Becca Heller , head of the International Refugee Assistant Project , said Monday . " Each new version is just an updated version of the previous ones , and they all have one goal to keep out people primarily from majorityMuslim countries . Any supposed national security goals are not based in fact . Instead , millions of people will be harmed indefinitely , including American families and institutions . We will not rest until this cruel and senseless policy is blocked for good "

(b) Huffington Post

Figure 1 An Example of Bias Word Detection

ferent nuance by comparing the news articles in  $D$  based on the relative features. Figure 1 shows an example of biasing word detection. The two news articles are reporting federal judge's decision on Trump's executive order referred to "Trump travel van". Here, the pure IDF scores are used for relative features to examine biasing words. The pure IDF-based feature scores are categorized into three scales, such as top 50% to top 30%, top 30% to top 10% and top 10%, highlighted by yellow, pink and red, respectively. It means that red colored words have more potential to be bias words. Topically important terms in the news article are colored with blue It can be identified by TF-IDF scores for each news article.

The paragraph level aggregates biasing scores of detected words for each paragraph of  $d$ . If the aggregated score of a paragraph exceeds a certain threshold value, the paragraph is considered as biasing candidate paragraph. From this step, we check which paragraphs in  $d$  are more likely to show different meaning from those of the news article set  $D$ .

Finally, the document level summarizes biasing capacity for the target news article  $d$  by using the aggregated scores of the paragraph level. Based on the summarized value, we determine whether the target news  $d$  is biased or not and how much it is biased.

## 5. Contrasting News Bias

To support the news reader' s understanding of news bias, we suggest a news contrasting system based on the bias detection process. The system contrasts news articles in the document level, the paragraph level and the word level. Figure 2 describes the overview of the system interface.

The bottom of Figure 2 shows how the news articles are distributed and where the user-selected article is located. By using this news map, the user can check the aspect skewness of a specific news event in outline. In the document level, if the user selects a document  $d$  about the news event  $e$ , the system automatically dis-

plays a news map which shows the distribution of the news articles in terms of their aspects and contents. Then it marks where the target document  $d$  and its counterpart article are located in the map by comparing all document in the news set  $D$  and displays it in parallel by using the counterpart choosing mechanism.

In the paragraph level, on the user-selected news article, bias candidate paragraphs are highlighted as the user scrolls down the article. Also the highlighted paragraphs are matched with those of the counterpart article. When the user scrolls down the news document, if there is a bias inducing paragraph, the system highlights that part and also finds contrasting paragraph which contains most opposing meaning, then, displays it in the opposite site.

For this implementation, it needs to consider the selecting method of biasing paragraph and its counterpart and synchronizing method of the paragraph pair automatically.

In the paragraph level, on the user-selected news article, bias candidate paragraphs are highlighted as the user scrolls down the article. Also the highlighted paragraphs are matched with those of the counterpart article. When the user scroll down the news document, if there is a bias inducing paragraph, the system highlights that part and also finds contrasting paragraph which contains most opposing meaning, then, displays it in the opposing site. The red boxes in the Figure 2 illustrates this process. For this implementation, it needs to consider the selecting method of biasing paragraph and its counterpart and synchronizing method of the paragraph pair automatically.

In the word level, the system highlights the detected bias inducing words by the three different degrees, such as high, middle and low, according to their biasing scores on the user-selected paragraph. In addition, the system shows the matched words which have contrast meanings in the opposite part based on the detected candidates for the bias words.

The main benefit of the system all that the bias can be displayed three different level and compared with the counterpart at the same

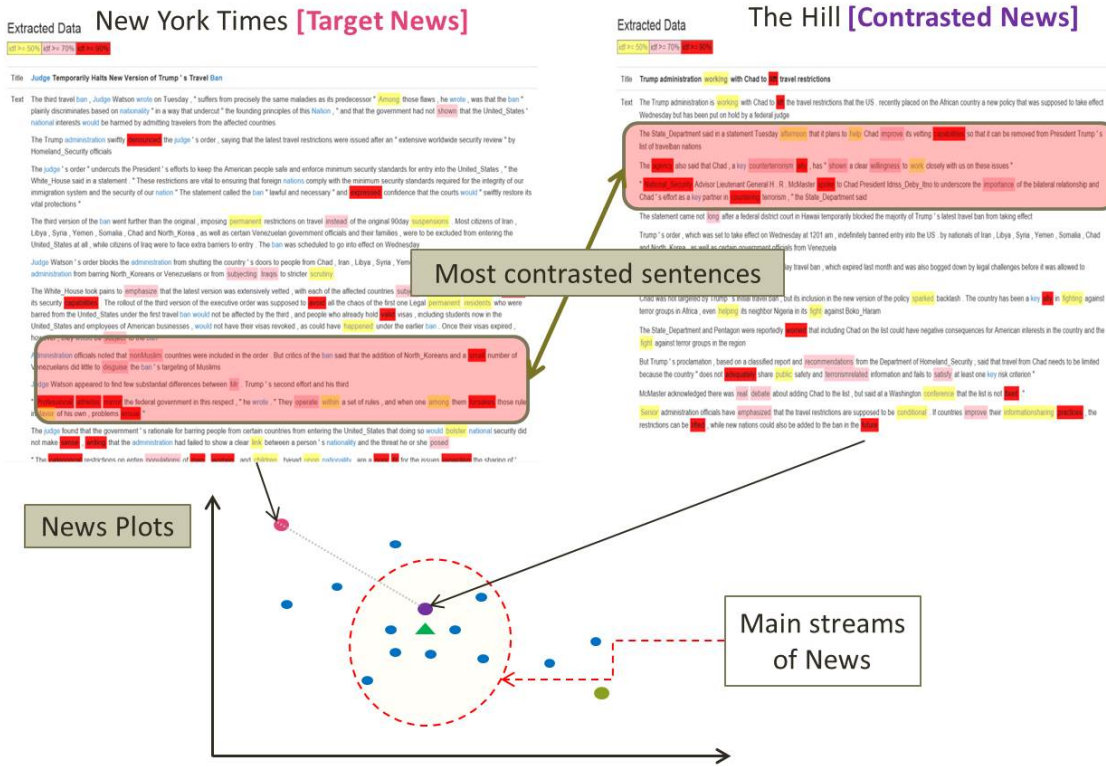


Figure 2 System Outlook for Bias Detection and Contrast

time so that user can understand the way of bias injection to the news articles caused by word choices. Because people generally perceive an object by comparing of others, the system could support better awareness of news bias as well as the way of word choices to generate news bias.

Our next step is to implement the detailed mechanism. Also, we will build ground truth for the bias words in the news articles and evaluate the performance of bias word detection by employing cloud resources. We anticipate our visualized interface can be used for this process.

## 6. Conclusions and Future Works

Detecting news bias is a challenging task in computer science as well as linguistics area because of its subtle and implicit characteristics. In this paper, we proposed relative features which adopt the concept of relativity to detect bias inducing words in news articles. We also presented two hierarchical approaches for news bias detection. Based on the bias detecting methods, we suggested a contrasting system for investigating news bias. We anticipate the system can support better awareness of news bias as well as the way of word choices to generate news bias.

In the future work, it needs to define the bias detection mechanism in detail. Specifically, it is about how to calculate partial difference but overall similarity. We will construct labeled dataset to evaluate how well the bias detecting process based on relative features captures bias in the newspaper by utilizing human resources. We will also enrich our features by applying subcategories from Verb-

Net [22], Framenet [20], and LIWC sources [17] suitable for news bias.

## References

- [1] Héctor Martínez Alonso, Amaury Delamaire, and Benoit Sagot. Annotating omission in statement pairs. *LAW XI 2017*, pages 41–45, 2017.
- [2] Jisun An, Meeyoung Cha, Krishna P Gummadi, Jon Crowcroft, and Daniele Quercia. Visualizing media bias through twitter. In *Sixth International AAI Conference on Weblogs and Social Media*, 2012.
- [3] Stefano Baccianella, Andrea Esuli, and Fabrizio Sebastiani. Sentimentnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining. In *LREC*, volume 10, pages 2200–2204, 2010.
- [4] Eric Baumer, Elisha Elovic, Ying Qin, Francesca Polletta, and Geri Gay. Testing and comparing computational approaches for identifying the language of framing in political news. In *Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 1472–1482, 2015.
- [5] Felix Hamborg, Norman Meuschke, and Bela Gipp. Matrix-based news aggregation: Exploring different news perspectives. In *Digital Libraries (JCDL), 2017 ACM/IEEE Joint Conference on*, pages 1–10. IEEE, 2017.
- [6] Joan B Hooper. On assertive predicates. *Syntax and Semantics*, pages 91–124, 1975.
- [7] Ken Hyland. *Metadiscourse: Exploring Interaction in Writing*. Continuum, London and New York, 2005.
- [8] Lauri Karttunen. Implicative verbs. *Language*, pages 340–358, 1971.
- [9] Bing Liu, Mingqing Hu, and Junsheng Cheng. Opinion observer: analyzing and comparing opinions on the web. In *Proceedings of the 14th international conference on World Wide Web*, pages 342–351. ACM, 2005.
- [10] Qiang Ma and Masatoshi Yoshikawa. Topic and viewpoint extraction for diversity and bias analysis of news contents. In *Advances in Data and Web Management*, pages 150–161. Springer, 2009.

- [11] Akiyo Nadamoto, Ma Qiang, and Katsumi Tanaka. Concurrent browsing of bilingual web sites by content-synchronization and difference-detection. In *Web Information Systems Engineering, 2003. WISE 2003. Proceedings of the Fourth International Conference on*, pages 189–199. IEEE, 2003.
- [12] Akiyo Nadamoto and Katsumi Tanaka. A comparative web browser (cwb) for browsing and comparing web pages. In *Proceedings of the 12th international conference on World Wide Web*, pages 727–735. ACM, 2003.
- [13] Tatsuya Ogawa, Masatoshi YOSHIKAWA, et al. News bias analysis based on stakeholder mining. *IEICE transactions on information and systems*, 94(3):578–586, 2011.
- [14] Eli Pariser. *The filter bubble: What the Internet is hiding from you*. Penguin UK, 2011.
- [15] Souneil Park, Seungwoo Kang, Sangyoung Chung, and Junehwa Song. Newscube: delivering multiple aspects of news to mitigate media bias. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 443–452. ACM, 2009.
- [16] Souneil Park, Minsam Ko, Jungwoo Kim, Ying Liu, and Junehwa Song. The politics of comments: predicting political orientation of news stories with commenters’ sentiment patterns. In *Proceedings of the ACM 2011 conference on Computer supported cooperative work*, pages 113–122. ACM, 2011.
- [17] James W Pennebaker, Ryan L Boyd, Kayla Jordan, and Kate Blackburn. The development and psychometric properties of liwc2015. Technical report, 2015.
- [18] Marta Recasens, Cristian Danescu-Niculescu-Mizil, and Dan Jurafsky. Linguistic models for analyzing and detecting biased language. In *ACL (1)*, pages 1650–1659, 2013.
- [19] Ellen Riloff and Janyce Wiebe. Learning extraction patterns for subjective expressions. In *Proceedings of the 2003 conference on Empirical methods in natural language processing*, pages 105–112. Association for Computational Linguistics, 2003.
- [20] Josef Ruppenhofer, Michael Ellsworth, Miriam RL Petruck, Christopher R Johnson, and Jan Scheffczyk. *Framenet ii: Extended theory and practice*, 2006.
- [21] Karen Sanders. *Ethics and journalism*. Sage, 2003.
- [22] Karin Kipper Schuler. *Verbnet: A broad-coverage, comprehensive verb lexicon*. 2005.