HIGH SCHOOL EDITION
Journal of Student Research

Volume 12 Issue 1 (2023)

# Analyzing the Extent to Which Gender Bias Exists in News Articles Using Natural Language Processing Techniques

Nihita Guda

Broad Run High School, Ashburn, VA, USA

## ABSTRACT

Prior studies have shown the existence of gender bias in job postings, performance reviews, and letters of recommendation. However, very little research has been done on the presence of gender biases in mainstream news sources and how they vary across publications. Human editing, given the rapid pace of news dissemination, is not effective enough to address biases. Even computer programs that parse the news articles for specific words and references still fall short of identifying and detecting the undertones and implicit references, which is why sophisticated techniques like Artificial Intelligence (AI) are necessary. In this study, I used Natural Language Processing (NLP) methods, a series of Python-programs to further analyze how biases vary in new information, along the metrics of type, variety, and intensity. I used over 500,000 news articles from 15 publications, spanning over 4 years to build and train my algorithm. Using Word2Vec, a popular NLP method, I was able to conclude that more right leaning publications are more likely to exhibit misogynistic content that is biased against women. However, the method fell short of identifying many forms of objectification like Benevolent Sexism. Similarly, using VADER, a python-code of sentiment analysis tool, I was able to determine that mere metrics of positive, negative, and neutral sentiment are not sufficient to detect occurrences of gender bias. To gauge the breadth of sexist language effectively, I used the LIWC text analysis program which calculates the percentage of words in a given text that fall into one or more of over 80 linguistic, psychological and topical categories indicating various social, cognitive, and affective processes. As a result, with statistical evidence my study was able to conclude the presence of implicit gender bias occurs all across publications but is more prevalent in right-leaning publications.

## Background Literature

The sheer quantity of news sources, the vast number of publications and methods of accessing them, and the increased polarization of opinions make it difficult to assess the reliability and validity of a piece of information. Deciding which publications and news to read is very important, because they have the potential to instill implicit biases in readers on grounds of race, gender identity, sexuality, socioeconomic status, education level, or other differences that may not even be realized. This is due to the concept of implicit biases because the type of media a person consumes affects not only the way they think, but the extent to which they think in that particular way. Considering that media is being consumed at a faster rate than ever before, it is especially important to shed light on the biases present in these articles.

A pervasive issue in society is the sexist and misogynistic language used against women. Women often fall victim to this language in the media (eg: social media like Reddit[1]) for the purpose of furthering an argument, and this has the potential to transform into implicit biases that the reader holds. This trajectory shows that

[1] Exploring Misogyny across the Manosphere in Reddit, http://oro.open.ac.uk/61128/1/WebScience139.pdf

it is important to consume information cautiously and critically. Through the use of various Artificial Intelligence (AI) techniques, it is possible to shed light to the extent to which gender bias is present in certain texts. This research has been done on performance reviews[2], letters of recommendation[3], and job postings[4]; however, very little research has been done on the presence of gender bias in popular media. By being able to analyze the extent to which gender bias exists in various publications, its readers will be able to possess the ability to consume media more carefully and cautiously.

What is gender bias, and what does it entail in the context of the media? To explore this concept further, we must first define "bias". The general definition of bias is the tendency to unfairly favor or hold prejudice against a thing, person, or group. This has many factors of influence, including environment, culture, personal experience, etc. Sometimes, a person is conscious of their biases – this is known as explicit bias. More often, people hold unconscious biases, known as implicit bias. This can cause someone to unintentionally discriminate or hold prejudices against particular groups of people or ideologies or can cause them to typecast people or groups – even if this social categorization is not factually accurate or reflective of conscious views. For example, if a teacher is unconsciously biased against girls in sports, they might always pick a boy to be the team captain, even if they may outwardly feel that girls and boys are equally as talented and enthusiastic about sports. An unconscious bias like this may be the result of gender stereotypes that imply that boys like sports while girls prefer art, in addition to societal norms that place more value on male competitive sports than their female equivalent. It is important to remember that individuals acting upon their biases does not necessarily indicate their malicious intent. Instead, biases form over our lifetimes, and we must fully recognize them and their sources in order to tackle them.

Media bias can affect the selection of events and stories that get published, the perspective from which they're written, and the language chosen to tell them. In most countries, media bias is thought to exhibit partisanship, meaning it either favors liberal or conservative politics. Media bias can also go so far as to completely reflect the ideals of the governing body, where the bias essentially becomes propaganda. Although a lot of media contains unconscious bias, in which journalists may be met with practical limitations to neutrality, media also regularly contains explicit bias. This is where media outlets deliberately try to paint a certain image of an event, group, or individual to achieve their desired outcome. This outcome may be politically fueled, or could be an attempt to appeal to a certain audience for monetary gain.

Gender bias exists in many forms, and it is important to be able to recognize and understand these forms in order to identify them. Sexual objectification is when a person's body – usually a woman's – is singled out and separated from her to be viewed as a mere physical object of sexual desire, according to the American Psychological Association (ASA). Benevolent sexism (a term coined by Glick, P., & Fiske, S. T.[5]) is an attempt to evaluate gender in a particularly positive way that is fully subjective to the evaluator, but is instead detrimental to the ideals of gender equality at large. For example, believing that women must be protected by men and must remain at home to bear children is a benevolently sexist outlook. Hostile sexism is an unashamedly brash reflection of negative stereotypes and ideologies related to gender. For example, the idea that women are inferior to men in mathematics is a hostile outlook and a claim with no valid evidence.

In order to understand the angles of gender bias in media, it is important to analyze the emotion and sentiment of the texts at hand, in addition to understanding the tone and usage of particular words or phrases

[2] How Gender Bias Corrupts Performance Reviews, and What to do about it, https://hbr.org/2017/04/how-gender-bias-corrupts-performance-reviews-and-what-to-do-about-it

[3] Defying a Gender Narrative, https://www.insidehighered.com/news/2022/03/09/study-challenges-gender-bias-letters-recommendation

[4] Evidence That Gendered Wording in Job Advertisements Exists and Sustains Gender Inequality, https://ideas.wharton.upenn.edu/wp-content/uploads/2018/07/Gaucher-Friesen-Kay-2011.pdf

[5] https://blogs.scientificamerican.com/psysociety/benevolent-sexism/

normally associated with gender-related bias in their respective contexts. It is important to note that misogyny and sexism can occur in many different ways: sexualization/objectification, benevolent sexism, and hostile sexism, and these occurrences appear with different keywords and sentiments in the media – meaning we can track them.

## Data

The dataset used to conduct this research is called "All the news" and was acquired from Kaggle's[6] database. This dataset consists of 143,000 news articles from all different sides of the political spectrum and comprises articles from these publications mainly from the beginning of 2015 to July 2017. This range is significant because it includes the time period in which presidential elections were occurring between the candidates Donald Trump and Hillary Clinton. During this time, there occurred lots of aggression from news sources, rendering an opportunity to see just how much women will fall victim to misogyny as news articles vigorously push their political agendas. The publications in this dataset include the New York Times, Breitbart, CNN, Business Insider, The Atlantic, Fox News, Talking Points Memo, Buzzfeed News, National Review, New York Post, The Guardian, NPR, Reuters, Vox, and the Washington Post. The Distribution of the number of articles in the dataset is shown in Figure 1 below. There is evidently a greater number of articles from certain publications such as Breitbart, New York Post, and Washington post and a smaller amount from Vox and Fox news. This suggests that the results will be more accurate for those publications with a higher number of articles which is important to keep in mind as we continue with the study.

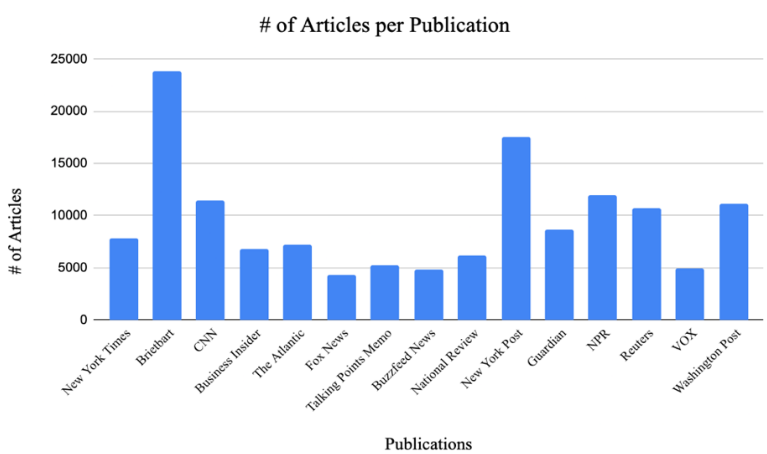


**Figure 1.** Distribution of Publications

[6] https://www.kaggle.com/datasets/snapcrack/all-the-news

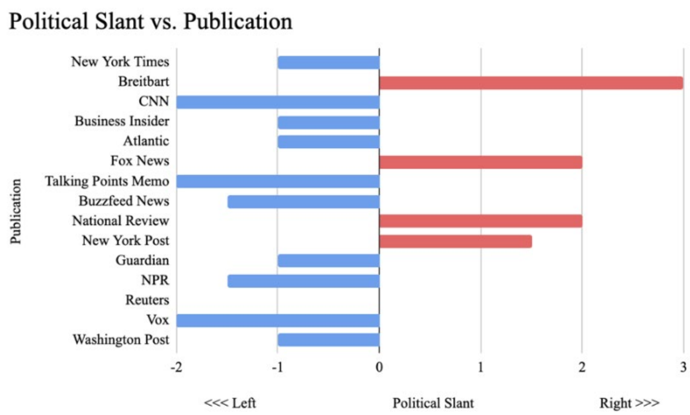


**Figure 2.** Political Slant of Publications

## Methods

One of the key steps in Machine Learning process is the need to pre-process large datasets into measurable pieces of data that can be used for analysis and training an AI algorithm. Through employing the Pandas and Natural Language Toolkit (NLTK) libraries, I was able to iterate through my dataset consisting of over 500,000 news articles in order to prep it for testing. The NLTK library allowed me to preprocess my data by stemming, lemmatizing, and tokenizing for easier analysis by helping map multiple words to a common root word. Once preprocessing has occurred, more explicit data mining tactics are used such as Word2Vec (W2V), Sentiment Analysis (VADER), and LIWC to detect bias at a deeper level.

### Female and Male Sentence Classifier

To effectively analyze how bias effects women specifically, there needs to exist some way to extract and distinguish between male and female sentences. In order to do this, I used spaCy[7] which is a software library for advanced natural language processing that has a built-in parts of speech tagger. Using this, I extracted the subject and object of a sentence. If either the subject or object contained these words: ['he', 'him', 'his', 'men', 'husband', 'man', 'son', 'Mr.'] it was classified as a male sentence, and if the subject or object contained these words ['she', 'her', 'hers', 'women', 'wife', 'Ms.', 'Mrs.', 'daughter', 'woman'] it was classified as a female sentence. For each publication, each article was tokenized by sentence to effectively iterate through and create a collection of all the "male" and "female" sentences per article to perform further analyses. For the Male sentences, the method's Precision is 0.91 and Recall is 0.85. For the Female sentences, the method's Precision is 0.89 and Recall is 0.86.

### Word2Vec (W2V)

Word2vec is a technique for natural language processing published in 2013. The word2vec algorithm uses a neural network model to learn word associations from a large corpus of text. Once trained, such a model can detect synonymous words or suggest additional words for a partial sentence. As the name implies, word2vec represents each distinct word with a particular list of numbers called a vector. The vectors are chosen carefully such that a simple mathematical function (the cosine similarity between the vectors) indicates the level of semantic similarity between the words represented by those vectors.

[7] https://github.com/explosion/spaCy

HIGH SCHOOL EDITION
Journal of Student Research
Volume 12 Issue 1 (2023)

Cosine similarity is the normalized dot product of 2 vectors and this ratio defines the angle between them. Those ratios can be used to establish a word's association with other words (e.g. "man" is to "boy" what "woman" is to "girl"), or cluster documents and classify them by topic. For example, the cosine similarity between the word "man" and the word "king" is the number x. Word2vec has the ability to replicate the relationship (x) between "man" and "king" when given another word, say "woman". A sort of analogical relationship can be established when inputting most_similar (positive=["king", "he"],negative=["she])) for which the output is "queen".

### Sentiment Analysis (VADER)

VADER (Valence Aware Dictionary and sEntiment Reasoner) is a lexicon and rule-based sentiment analysis tool that is specifically attuned to sentiments expressed in social media. It is fully open-sourced under the MIT License[8].

The Python code for the rule-based sentiment analysis engine implements the grammatical and syntactical rules by incorporating empirically derived quantifications for the impact of each rule on the perceived intensity of sentiment in sentence-level text. It incorporates word-order sensitive relationships between terms. For example, degree modifiers (also called intensifiers, booster words, or degree adverbs) impact sentiment intensity by either increasing or decreasing the intensity. Consider these examples: (a)"The service here is extremely good" (b) "The service here is good" (c) "The service here is marginally good" Using a degree modifier increases the positive sentiment intensity of example (a) Likewise, example (c) reduces the perceived sentiment intensity using the word "marginally". I applied and compared the sentiment scores using the Female and Male Sentence Classifier to distinguish between the male and female sentences.

### LIWC (Linguistic Inquiry and Word Count)

To gauge the breadth of sexist language effectively, I utilized the LIWC software that is a text analysis program that calculates the percentage of words in a given text that fall into one or more of over 80 linguistic, psychological and topical categories indicating various social, cognitive, and affective processes. This will allow us to pinpoint the type of sexist language occurring in order to see how the compare across publications. I applied and compared the LIWC scores using the Female and Male Sentence Classifier to distinguish between the male and female sentences.

## Results & Discussion

### Word2Vec (W2V)

**Table 1.** Analogy Output Sample

| Breitbart | National Review | New York Post | Business Insider | VOX | NPR |
|---|---|---|---|---|---|
| sexy | boring | feminine | unforced | visual | realistic |
| creepy | skiing | romantic | competent | authentic | readily |
| masculine | entertaining | creepy | authentic | quirky | achievable |
| scientifically | slayers | fabulous | frustrating | abstract | subjective |
| fluid | charming | masculine | tremors | poignant | straightforward |
| catchy | sexy | retro | interfaces | storytelling | neat |

[8] https://github.com/cjhutto/vaderSentiment/blob/master/LICENSE.txt

| | | | | | |
|---|---|---|---|---|---|
| unpleasant | dumb | whimsical | novel | witty | tricky |
| insensitive | swarmed | beading | tweaks | provocative | perverse |
| reactionary | misinformed | erotic | robes | graphic | confusing |
| sinister | traumatized | curvy | genuinely | relatable | manageable |

man : clever :: woman : ? *(window=5, min=2)*

The W2V results were able to exhibit analogical relationships between certain words. The most telling occurrence of gender bias in these results were illustrated in the analogy *"man is to clever as woman is to …"*. If the news had displayed men and women in the exact same light, the results of this analogy would most likely include words like "clever" and similarly synonymous words. However, across all publications, this was rarely the case.

In the above table, three more left leaning publications (Business Insider, VOX, and NPR) and three more right leaning publications (Breitbart, National Review, and New York Post) were chosen to illustrate the difference in results. In the left leaning publications, the results mostly included words that denoted women's emotional responses, aptitude, and utility. Only in very few cases were words that denoted a level of smartness displayed. Furthermore, the right leaning publications display objectifying benevolent and hostile sexism to a much greater extent than the left leaning publications. This is evidence that right leaning publications take more harsh attacks at women than the left leaning publications, thus leaving their audience more susceptible to internalizing these biases.

## Sentiment Analysis (VADER)

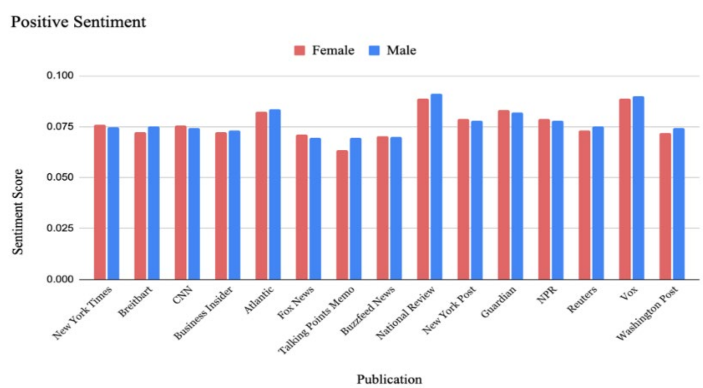


**Figure 3.** Positive Sentiment Analysis

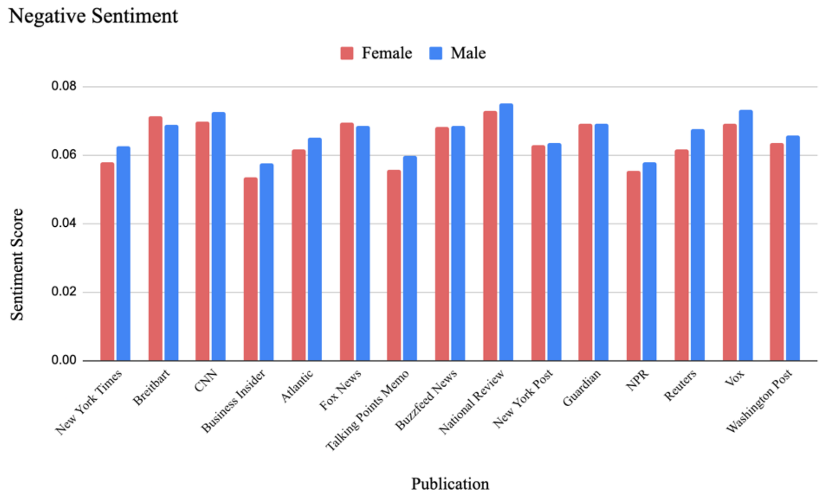


**Figure 4.** Negative Sentiment Analysis

The VADER tool looked at the *Positive*, *Negative*, *Neutral*, and *Compound* sentiments of sentences, in which the subject and/or object was female or male. This technology indicated how sentiment level differs between the two genders. The results, as shown in Figure 3 and Figure 4, show no statistically significant evidence that gender bias exists, or that women are depicted in a more positive or negative light compared to their male counterparts. This, however, contradicts our findings from above, which indicate how women are depicted as less "intelligent" than men. This data is indicative of the non-dichotomized nature of gender bias and confirms that it cannot be analyzed from the shortsighted perspective of only a positive, negative, or neutral sentiment. Different forms, angles, and subtleties of gender bias exist, and these can only effectively be analyzed through a multidimensional lens that considers the varying ways in which gender bias can be exhibited. Relevancy of the angles of gender bias – benevolent and hostile sexism – come into use.

## LIWC (Linguistic Inquiry and Word Count)

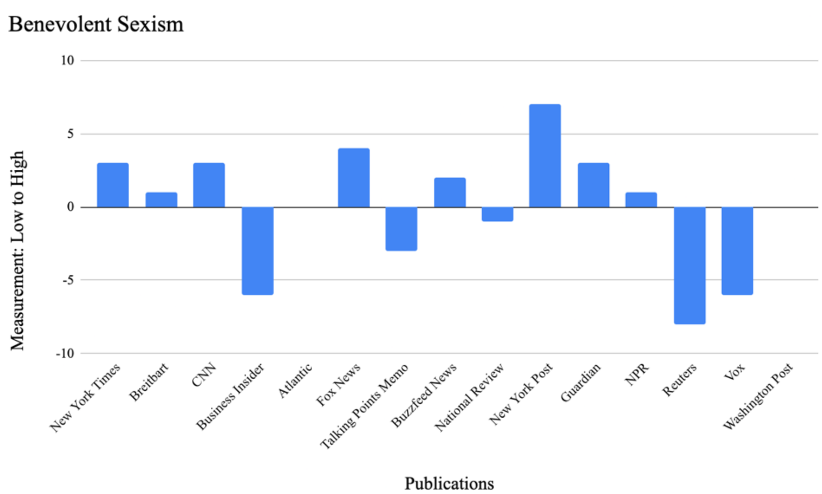


**Figure 5.** Benevolent Sexism

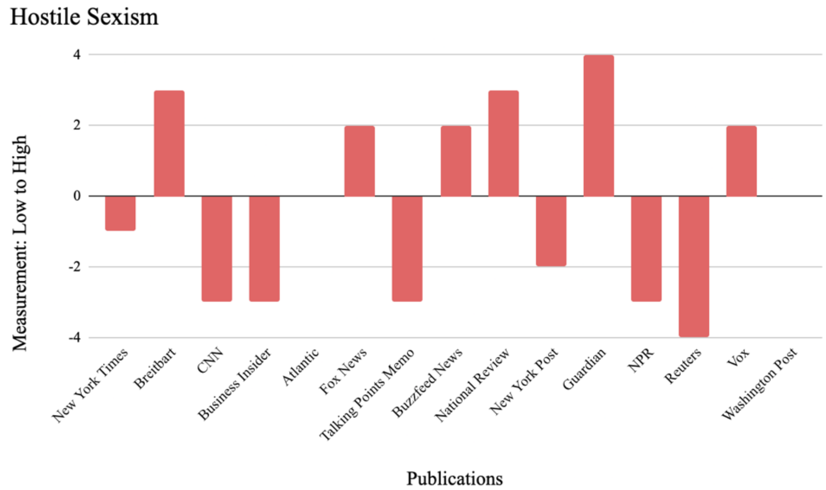


**Figure 6.** Hostile Sexism

The LIWC allows many different *dimensions* of sentiment to be analyzed in order to more effectively gauge gender bias, which can be easily disguised in the media. To effectively identify and capture instances of it, the LIWC has the ability to look at many tones, emotions, and drives, throughout texts. With this ability, the LIWC can more easily observe occurrences of gender bias across various publications. The vast majority of publications exhibit some sort of bias, but what kind of bias is it and how does it correlate to the ideological slant of the publications?

Benevolent and Hostile sexism had different LIWC metrics that measured them. For example, benevolent had “social”, “polite”, “home”, “family”, etc. while hostile had “negative”, “anxiety”, “anger”, “swear”, etc. In order to measure how they differ across publication, for each metric, I extracted the three publications that had the highest and lowest metric score. At the end, going by publication, I subtracted the number of lowest occurrences from the number of highest occurrences. Say for example, Buzzfeed news was one of the highest scoring out of all the publications for three different hostile sexism metrics (“anxiety”, “anger”, and “negative” for example), and was also one of the lowest scoring for one hostile sexism metric (“swear” for example). To calculate the measure of the publication, I would subtract one from three to determine the “net” hostile sexism that publication exhibited. That is how the above tables have been calculated.

As shown in Figure 5 both right and left leaning publications exhibit benevolent sexism, but the right leaning publications exhibit it to a slightly greater extent. As shown in Figure 6, the more right-leaning publications tend to exhibit hostile sexism to a much greater degree than left leaning publications.

## Conclusion

This research paradigm was developed to investigate the extent to which gender bias exists in mainstream news sources. I hypothesized the left leaning and “factual” publications may exhibit a lesser extent of gender bias, and this was supported by my data. Furthermore, the types of biases exhibited in these publications differed from one another –both right and left leaning publications exhibited benevolent sexism with the right leaning publications exhibiting it to a slightly greater extent. On the other hand, the more right-leaning publications tend to exhibit hostile sexism to a significantly greater degree than left leaning publications. Publications from across the political spectrum displayed a moderate amount of benevolent sexism and discussed men as strong, confident, and dominant to a significantly greater extent than they did women. The dataset used for this study contained publications of varying sample sizes, which may have resulted in slightly more accurate results for publications with a greater number of articles and thus larger sample size. Furthermore, it was challenging to identify every occurrence of gender bias in a text because there is always some unaccounted-for ambiguity in whether a certain phrase is referring to a female or not. The method utilized for this paradigm took sentences with female subjects and objects and studied the sentiment within them for the VADER and LIWC tests, which is relatively comprehensive, but not all-encompassing of every single instance of gender bias. Thus, as per the data, this paper statistically suggests the existence of gender bias at a greater extent in right leaning publications in mainstream media.

## Acknowledgments

I would like to thank my advisor for the valuable insight provided to me on this topic.

## References

Allen, C. & Hospedales, T. (n.d.). *Analogies explained: Towards understanding word embeddings*. Retrieved June 1, 2022, from https://arxiv.org/pdf/1901.09813.pdf

Beri, A. (2020, May 27). *Sentimental analysis using VADER*. Medium. Retrieved May 31, 2022, from https://towardsdatascience.com/sentimental-analysis-using-vader-a3415fef7664

D'Ignazio, C. & Klein, L. F. (2020). *Data feminism*. The MIT Press.

Honnibal, M., & Montani, I. (2017). *spaCy 2: Natural language understanding with Bloom embeddings, convolutional neural networks and incremental parsing*.

Hutto, C.J. & Gilbert, Eric. (2015). VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text. Proceedings of the 8th International Conference on Weblogs and Social Media, ICWSM 2014.

Madera, J. M., Hebl, M. R., Dial, H., Martin, R. & Valian, V. (2018). Raising doubt in letters of recommendation for academia: Gender differences and their impact. *Journal of Business and Psychology*, 34(3), 287–303. https://doi.org/10.1007/s10869-018-9541-1

Mastari, L., Spruyt, B. & Siongers, J. (2019). Benevolent and hostile sexism in social spheres: The impact of parents, school and romance on Belgian adolescents' sexist attitudes. *Frontiers in Sociology*, 4(47). https://doi.org/10.3389/fsoc.2019.00047

Nittle, N. (2021, April 28). *The language of gender bias in performance reviews*. Stanford Graduate School of Business. Retrieved May 31, 2022, from https://www.gsb.stanford.edu/insights/language-gender-bias-performance-reviews

Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. New York University Press.

Salmen, A. & Dhont, K. (2020). Hostile and benevolent sexism: The differential roles of human supremacy beliefs, women's connection to nature, and the dehumanization of women. *Group Processes & Intergroup Relations*, 24(7), 1053–1076. https://doi.org/10.1177/1368430220920713

Vatsal. (2022, May 20). *Word2Vec explained*. Medium. Retrieved May 31, 2022, from https://towardsdatascience.com/word2vec-explained-49c52b4ccb71