Does Gender Matter in the News? Detecting and Examining Gender Bias in News Articles

Jamell Dacon*
daconjam@msu.edu
Michigan State University
East Lansing, Michigan, USA

Haochen Liu
liuhaoc1@msu.edu
Michigan State University
East Lansing, Michigan, USA

ABSTRACT

To attract unsuspecting readers, news article headlines and abstracts are often written with speculative sentences or clauses. Male dominance in the news is very evident, whereas females are seen as "eye candy" or "inferior", and are underrepresented and under-examined within the same news categories as their male counterparts. In this paper, we present an initial study on gender bias in news abstracts in two large English news datasets used for news recommendation and news classification. We perform three large-scale, yet effective text-analysis fairness measurements on 296,965 news abstracts. In particular, to our knowledge we construct two of the largest benchmark datasets of possessive (gender-specific and gender-neutral) nouns and attribute (career-related and family-related) words datasets[1] which we will release to foster both bias and fairness research aid in developing fair NLP models to eliminate the paradox of gender bias. Our studies demonstrate that *females* are immensely marginalized and suffer from socially-constructed biases in the news. This paper individually devises a methodology whereby news content can be analyzed on a large scale utilizing natural language processing (NLP) techniques from machine learning (ML) to discover both implicit and explicit gender biases.

CCS CONCEPTS

• **Computing methodologies** → **Natural language processing**; *Artificial intelligence*; • **Social and professional topics** → *Gender*.

KEYWORDS

Media bias, News bias, Gender bias

ACM Reference Format:
Jamell Dacon and Haochen Liu. 2021. Does Gender Matter in the News? Detecting and Examining Gender Bias in News Articles. In *Companion Proceedings of the Web Conference 2021 (WWW '21 Companion), April 19–23, 2021, Ljubljana, Slovenia.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3442442.3452325

1 INTRODUCTION

In recent years, there has been a growing popularity of online newspapers in comparison to traditional "printed" newspapers [38]. A benefit to online news is that news articles are constantly updating; furthermore, news titles and abstracts are regularly taken into consideration when recommending news to quickly attract users [11]. However, to attract the attention of users, rich textual information such as news titles and abstracts present various forms of media biases such as ideological bias (i.e., biased articles that attempt to promote a particular opinion on a topic), coverage bias (i.e., media coverage in regards to the visibility of topics or entities), selection bias, and presentation bias [16], thus contributing to the problem of *gender bias*. Since the 1950s, there have been studies on biased news reporting [36]. Media bias is both intentional as it reflects a conscious act and is sustained to present a systematic biased tendency[37]. Male dominance is well documented, and in news articles, men are always depicted as leaders while women are depicted as '*inferior*' or as '*eye candy*' [21]. Nevertheless, consumers of online news services are attracted to novelty and/or differences such as skin-color, ethnicity, gender identity, or sexual orientation, which creates an ingrained feeling of interest or curiosity that may result in chronic socially-constructed biases.

News articles are often written with speculative sentences or clauses to clinch a reader's attention [12], and thus, play a crucial role in shaping public and personal opinions on public affairs and political issues [16]. An example of explicit informational bias in gender-specific (*male* and *female*) job promotion news titles[2] is, "*Women who want to succeed at work should shut up - while men who want the same should keep talking, research says*", compared to, "*Men have been promoted 3 times more than women during the pandemic, study finds*". In this example, those titles present enough information about the news' *body* content; however, in some cases, titles may not have enough textual information. For example, "*Women in the workplace.*", whereas an abstract will possess a quick overview of the news article, therefore, containing sufficient information content to indicate the presence of gender bias. Although online news recommendations [11, 38] continuously provide novel news stories, the textual information demonstrates and constitutes socially-constructed biases. Women represent nearly half of the world's population, yet they are greatly under-examined and underrepresented in news stories [21]. Those who are considered to be newsworthy are politicians, CEOs, engineers, doctors, pilots, basketball players, and so on – are often men. When women are considered to be newsworthy they are often presented as sexual beings for their *bodies*, *motherhood*, and/or being *supportive wives* [19, 20]. In short, news media heavily influences gender roles in society by serving as a basis of stereotypes which results in the reinforcement of social inequalities. Therefore, conveying categorical

*Corresponding author
[1] https://github.com/daconjam/Detecting_Gender_Bias

[2] https://www.huffpost.com/entry/daily-mail-headlines-women-ridiculous_n_2192332

WWW '21 Companion, April 19–23, 2021, Ljubljana, Slovenia

ACM ISBN 978-1-4503-8313-4/21/04.
https://doi.org/10.1145/3442442.3452325

barriers, and thus, controlling ones' self-identity and determining ones' position in a hierarchical taxonomy.

Natural language processing (NLP) techniques and systems aim to learn from natural language data, and mitigating social biases becomes a compelling matter not only in machine learning (ML) but for social justice as well. Sentence and word embeddings are popular NLP tools that capture the semantic similarities of sentences and words which display human-like societal biases [2, 3, 18, 27], whereas text classification [40] also know as *text tagging* is a computational process of categorizing texts into groups. Several NLP text classifiers can assign a set of predefined *tags* by automatically analyzing texts based on their textual information. Previous existing works have taken different approaches to address the issue of gender bias by detecting the *male/female* ratio of images [19, 20], measuring fairness in dialogues systems [10, 23], language modeling [4], machine translation [7], and coreference resolution [42].

In this work, we conduct an innovative study of bias issues in gender representation in news abstracts in two large English news datasets i.e., MIND Dataset[3] [38], and a News Category Dataset[4] [31] which are two large scale high quality news datasets constructed for news recommendations and news classification. Our goals are to detect and examine the phenomenon of implicit (i.e., bias that is implied and not stated directly) and explicit (bias that is plainly stated) gender bias in the abstracts of news articles where information about gender related stories to gain a sense of understanding of the gender representation in the news by examining the relationships between social hierarchies and news content. Our motivation is to identify how several forms of bias such as coverage bias, selection bias, and presentation bias contribute to the problem of gender bias. As gender fairness in news articles is an important problem, we analyze representational harms such as ideological bias which inseminates adverse generalizations about women.

(1) We construct two large benchmark datasets: (1) possessive (gender-specific and gender-neutral) nouns dataset and (2) attribute (career-related and family-related) words dataset to study gender bias, and we will release them to foster both bias and fairness research;
(2) We systematically conduct large scale analyses of each news corpora to detect and examine gender biases in distribution, content, and labeling and word choice;
(3) We demonstrate that there exist conclusive socially-constructed biases in regards to gender by introducing a series of measurements to better understand gender representation in news articles quantitatively and qualitatively.

2 RELATED WORKS

The elimination of gender discrimination is an important issue that contemporary society is facing. Gender bias is reflected in various behaviors of people, among which language is one of the most powerful means to express sexism [22, 30]. Existing works analyze gender bias in language of different fields. Madera et al. [26] discuss the gender stereotypes reflected in job evaluation languages such as letters of recommendation for academic positions. Gaucher et al. [14] analyze the gendered wordings used in job advertisements and discuss how they reflect gender inequality. In the field of education, gender bias in high school textbooks [1] and computer science education materials [28] are studied. Menegatti and Rubini [30] investigate gender bias in general language usages. The authors discuss two types of gender bias in languages: the unfair lexical choices caused by gender stereotypes and the sexism embedded in language structures, including grammatical and syntactical rules. The authors emphasize the beneficial effects of gender-fair linguistic expressions and suggest to mitigate gender bias by using them. Recently, Pinto et al. [33] extend this line of research to the field of law. The authors study the gender bias reflected in the languages of court decisions. As a pioneering work, we investigate the gender bias in news languages in this paper to promote gender equality in the field of journalism.

Man-made text data are widely used to train machine learning models for various NLP tasks. Learning from human behaviors, NLP models have been proven to inherit the prejudices from humans [25, 29]. Existing works attempt to address the issue of fairness in various NLP tasks such as text classification [5, 32, 39], word embedding [2, 15, 43], coreference resolution [35, 42], language modeling [4], machine translation [13], semantic role labeling [41], dialogue generation [23, 24], etc. In this paper, we are committed to a better understanding of gender bias in news texts, thus contributing to building fair NLP models trained on such data, such as news recommender systems, news classifiers, and fake news detection models, etc.

3 DATASETS

We first collect two English news datasets [9], i.e., MIND Dataset (MIND) and a News Category Dataset (NCD). In our corpus, we retrieved 363,385 news articles, thus 363,385 news titles. As previously mentioned in Section 1, some titles do not present enough informational content about an article's body content to attract users, hence the notion to analyze the abstracts of each news articles. We later extract a total of 296,965 news abstracts from 363,385 news articles. Following this, inspired by [23] we develop two large word datasets (1) Possessive nouns dataset: a large benchmark gender-specific possessive nouns dataset containing a total of 465 non-offensive masculine and feminine gender possessive nouns (see Appendix A.1 and A.2), and (2) Attribute words dataset: a large benchmark gender-specific and gender neutral dataset containing a total of 357 masculine, feminine and neutral career-related and family-related words (see Appendix B.1 and B.2). We then conduct the three experiments to detect and examine the bias across the two news datasets. We will now detail the two news recommendation and news classification datasets as follows:

- **MIND**: The MIND dataset was collected from the Microsoft News[5] website. Wu et al. [38] randomly sampled news for 6 weeks from October 12th to November 22th, 2019 to create two datasets i.e., MIND and MIND-small both totaling in 161,013 news articles. Each news article contains a news ID, a category label, a title, and a body (URL); however, not every article contains an abstract resulting in 96,112 abstracts. We used the training set (largest set of news articles) since both the validation and test sets are assumed to be subsets of

[3]https://msnews.github.io/
[4]https://www.kaggle.com/rmisra/news-category-dataset
[5]https://microsoftnews.msn.com/

Dataset	Abstracts	Category	M	F
MIND	96,112	18	22,760	6,817
NCD	200,853	41	21,250	15,856

Table 1: Gender distribution test on the news datasets.

	MIND		NCD	
Career Words	# Man	# Woman	# Man	# Woman
Spokes	192	121	112	42
Congress	191	49	94	25
Chair	225	20	102	5
Business	66	3	31	4

Table 2: Illustration of four intersecting career words (prefixes) across the two datasets for *females* compared to their respective *male* counter parts. The results are reported in terms of no. of gender-specific career words mentioned in each dataset per gender with their corresponding *Woman/Man* suffixes.

the training set. MIND is created to serve as a new news recommendation benchmark dataset.

– **NCD**: The NCD dataset [31] was collected from Huffpost[6]. The news articles were sampled from news headlines from the year 2012 to 2018 totaling in 202,372 news articles. Each news article contains a category label, headline, authors, link, and date; however, not every article contains a short description (abstract) resulting in 200,853 abstracts. NCD serves as a news classification and recommendation benchmark dataset.

4 BIAS IN GENDER DISTRIBUTION

In this section, we explore the gender distribution in news abstracts across the two datasets to determine the presence of category bias and occupational bias by identifying words in our possessive nouns and attribute words dataset (see Appendix A and B).

4.1 Gender Distribution

Gender distribution refers to the *diversity* in the abstracts of each news article. The distribution is a simple, yet key measurement of equality in the number of males to females in each news dataset. Given that an abstract contains one or more sentences or clauses consisting of gender identity terms, the intuition is to classify a **sex**, *i.e., male* (**M**) or *female* (**F**), otherwise *neutral* **N** for each abstract. In turn, this quantification refers to the proportion of the number of *males* to *females* in each news category. Hence, we label each news abstract with one of three possible labels, (1) **M**: if the abstract contains more masculine possessive nouns, (2) **F**: if the abstract contains more feminine possessive nouns, and lastly, (3) **N**: if the abstract contains none or the same number of masculine and feminine possessive nouns. For neutral (**N**) cases, we simultaneously disregard unisex gender nouns *e.g. baby, child, employee, worker, etc.*, and people's names in abstracts as they can also be unisex, pet names, nicknames, or stage names for both males and females, *e.g. Max, Dylan, Jamie, Jordan, Blake, Taylor, etc,*.

[6]https://www.huffpost.com/

4.2 Experiment

In this measurement, we aim to investigate the gender distribution of *males* to *females* abstracts across the two news datasets. We first calculate the gender distribution in each dataset by parsing each sentence or clause of each abstract for gender identity terms to classify a sex, *i.e., male*: (**M**), *female*: (**F**), or *neutral*: **N**. As previously mentioned, to determine the sex of an abstract we label an abstract with one of three possible labels, **M** if the abstract contains more masculine possessive nouns; otherwise, **F** or **N** from a total of 465 masculine and feminine gender-specific and gender-neutral possessive nouns. Table 1 presents the results of the gender distribution test on the news datasets in terms of the total number of abstracts, categories per dataset, and the number of gender-tagged abstracts. One can observe that distribution results from MIND are quite distressing, as *female* abstracts are greatly underrepresented.

In NCD, *female* abstracts are not overly underrepresented; nonetheless, NCD possesses the largest number of categories and thus motivating the notion to investigate the category distribution in our now-labeled gender-tagged news abstracts. We examine the gender distribution across each category to identify if there exists a large proportion of gender biased topics e.g. *Politics*. As previously mentioned MIND was collected over a period of 6 weeks consisting of 18 categories; however NCD was collected over a period of 6 years consisting of 41 categories. We observe that **F** tagged abstracts are not underrepresented in NCD as females are over-represented in particular categories. We discover that the top 3 **F** tagged categories for NCD are *Style & Beauty, Parenting* and *Entertainment* which accounted for over 36% of the news reported in 41 categories, thus confirming that in the news articles collected over half of a decade that *females* are often presented in the news for motherhood and indeed often referred to for their physical characteristics.

Inspired by two recent works [23, 44], we construct an exhaustive list of career words to further explore the working class distribution to establish a sense of occupational mentions across the three datasets. This set is created from the the combination of occupational (career-related) words from Appendix A.1 and A.2 (see Appendix B.1). Unlike [23], we did not use generic gender-neutral *career words* such as engineer, dentist, lawyer, etc., but instead we use gender-specific career words such as ***policeman***, ***chairman***, ***spokesman***, etc., – and so on along with their respective *female* counterparts. Table 2 illustrates the top four intersecting career words for **F** compared to corresponding **M** gender-specific career words across the three datasets. Here, we see that within the news women suffer from several biases and are under-examined in regards to being acknowledged in the working class.

5 BIAS IN CONTENT

In this section, we investigate the occurrence frequency of career-related words and family-related words in news abstracts of different genders, where specific words reflect socially-constructed stereotypes of different genders, such as *females* being excessively associated with family words more than career words.

5.1 Attribute Words

In society, there are some socially-constructed stereotypes that heavily entail gender roles, *i.e.,* a specific gender is more anticipated with certain words. For example, society tends to identify *males* with career-related words and *females* with family-related words [6]. Words that influence gender roles in society, are known as *attribute words.* We use these attribute words to measure the fairness in each now-labeled gender-tagged news abstract by comparing the averages of attribute words that emerge in each abstract for each label. Inspired by the recent works [6, 23], we then proceed to construct a more exhaustive list of attribute words. In comparison, the career words list consists of both gender-specific and gender-neutral occupational (career-related) words, and family words list consists of both gender-specific and gender neutral family-related words to measure the fairness of each gender (see Appendix B.1 and B.2).

5.2 Experiment

In this measurement, we explore the average number of attribute words that appear in each gender-tagged abstract from a total of 357 masculine, feminine and neutral career-related and family-related words. As previously mentioned, females are excessively associated with family-related words more than career-related words, unlike men who are typically associated with career-related words. The bias measurement is straightforward, yet fundamental as it examines the occurrence frequency of career-related and family-related words in each gender-tagged news abstract to demonstrate the existence of socially-constructed stereotypes. To do so, we check both subsets of attribute words simultaneously. Table 3 presents the gender diversity which is simply the total percentage of gender-tagged abstracts across each dataset, and the average attribute words observed in each abstract across both news datasets.

One can observe that diversity results from MIND are poor as a result of *females* being greatly underrepresented in the news, however, NCD has diversity difference of 2.84% due to the over-representation in categories such as *Style & Beauty, Parenting* and *Entertainment* which acquired over a third of the NCD dataset, respectively. We observe that *males* are often associated with career-related words on average, and *females* are heavily and regularly associated with family-related words. These results are dismal as females are equally intelligent, thus these values should reflect similarity across both since both genders have the ability to advance in business.

6 BIAS IN WORDING

In this section we attempt to identify the influential terms i.e., the textual "centers" of the gender-tagged abstracts by applying two algorithms (1) Sentiment Analysis: to investigate the sentiment of an abstract's contextual information used to describe different genders across both news datasets, and (2) Centering Resonance Analysis: to discover the most central nouns that mostly contribute to the meaning of a document or corpora.

6.1 Sentiment Analysis

The sentiment of an abstract is crucial to examine if the opinions conveyed by the columnist are negative (Neg.), neutral (Neu.) or positive (Pos.). We apply the popular, well known sentiment analysis tool, VADER [17] to measure the sentiment of each news abstracts. VADER computes a normalized, weighted *compound* score of each word in a sentence by summing their valence scores between -1 (being extremely negative) and +1 (being extremely positive). As abstracts are usually one or more sentences, the abstracts are split into sentences to operate on a sentence level by employing an NLTK[7] toolkit sentence tokenizer. Therefore, if there are more negative sentences than positive and neutral sentences, we treat the abstract as negative. Otherwise neutral or positive. An example of a neutral abstract is, "*An auction of shares in Google, the web search engine which could be floated for as much as $36bn, takes place on Friday*". For oxymoronic news abstract cases where the number of positive and negative sentences are the same e.g., "*He finally got the promotion he so longed for! Unfortunately, his wife filed for divorce that same day.*", we treat the abstract as neutral. We simply use the compound score within the respective thresholds of positive, negative and neutral sentiments when considering the sentiment of an abstract containing only one sentence.

6.2 Centering Resonance Analysis

Corman et al. [8] contrast three objectives of computational text analysis as follows: Inference, Positioning and Representation [34]. The authors argue that a number of machine learning (ML) algorithms must be trained on a corpus before being applied, and that popular models such as Latent Semantic Analysis (LSA) or Latent Dirichlet Allocation (LDA) attempt to reduce a given text into a vector lying within the same semantic space. However, this encourages a narrow domain due to the quality of spatial construction and results in a loss of information. Therefore, there is a need for a representative method that can accomplish the three objectives. Centering Resonance Analysis (CRA) first proposed by Corman et al. [8] is a network word-based method that constructs a network representation of correlated words. This method exploits rich textual data and expresses the intention and meandering behaviors of authors (or columnists) [34]. CRA is able to determine textual "centers" without the use of dictionaries or being trained on a corpus i.e., identifying the most central nouns that mostly contribute to the meaning of a document or corpora.

[7] https://www.nltk.org

	Dataset			
	MIND		NCD	
	M	**F**	**M**	**F**
Diversity (%)	23.68	7.09	10.22	7.88
Avg. Career Words per Abstract	**0.1258**	0.0907	**0.0657**	0.0554
Avg. Family Words per Abstract	0.6406	**0.6954**	0.4431	**0.4723**

Table 3: The average number of the attribute words observed in each news abstract.

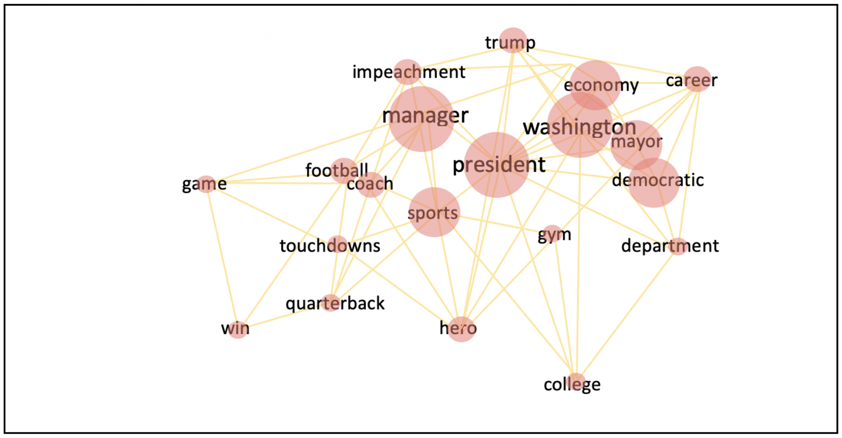

Figure 1: (a) The resulting CRA network for the top 20 nouns in the M tagged abstracts.

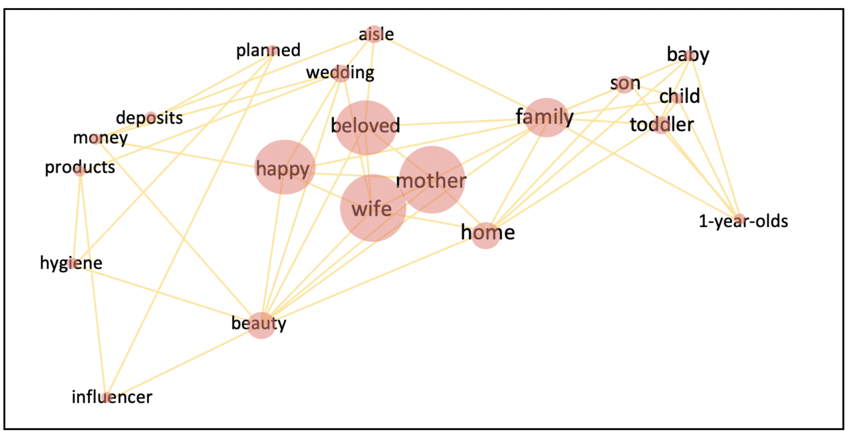

Figure 1: (b) The resulting CRA network for the top 20 nouns in the F tagged abstracts.

6.3 Experiment

In this measurement, we illustrate two representative text networks depicting the most central noun phrases for the combined gender-tagged abstracts. This bias measurement examines the compound noun phrases that are most prevalent for each gender. We measure the noun similarities between the two types of gender-tagged abstracts by combining both news datasets i.e. MIND and NCD and calculating the resonance of the now MIND+NCD dataset. We first apply the sentiment analysis tool to predict the sentiment of each sentence in each abstract for **M** and **F** tagged abstracts, i.e., when there are more positive sentences than negative and neutral sentences, abstracts are treated as positive; otherwise negative or neutral. We later aggregate the positive abstracts for both **M** and **F** as this implies the the most constructive attention for both *males* and *females*. We neglect the negative and neutral abstracts as we assume common noun phrases would be generic words used in adverse news articles. For example, ***killer, murderer***, – and so on.

After aggregating a total of 23,795 positive abstracts, we first remove stopwords as they capture little to no semantic information and more importantly reduces computational complexity. We then implement two algorithms: (1) an NLTK package for identifying compound noun phrases by tagging parts-of-speech (POS-tagger), and more specifically, it exploits a Penn Treebank Tagger to identify compound nouns and adjectives. However, since we are solely interested in nouns, we only examine them; and (2) NetworkX[8] for detecting and analyzing the centrality of networks, hence identifying the textual centers of each dataset. Figures **??**. (a) and (b) presents the CRA networks results of the most central and/or compound nouns found in each abstract for **M** and **F** tagged abstracts. Note that, a total of 33,871 *distinct* nouns are prominent in the structuring of the text. The network construction became computationally expensive and did not have much explainability due to its denseness. Therefore, we attempt to address the dense network issue by constructing CRA networks for the top 20 compound nouns (highest resonance scores) for both gender-tagged abstracts. Each graph illustrates the positive nouns that contribute the most to specific topics of the abstracts according to their respective textual centers.

The results are utterly disappointing as *females* (**F** tagged abstracts) are undoubtedly heavily associated with family words in comparison to males (**M** tagged abstracts) are often associated with political and occupational terms. The top 20 words females are densely associated with are *mother, wife, beloved, happy, home, wedding, family, beauty, son, child, toddler, baby, 1-year-olds, aisle, planned, deposits, money, products, hygiene* and *influencer*, respectively. While males are easily associated with *president, Washington, manager, economy, mayor, sports, democratic, impeachment, career, gym, trump, football, coach, hero, touchdowns, quarterback, game, win* and *college*, respectively. Thus, there exists a strict gender dichotomy of men and women. Even though women succeed at clichéd male tasks the nouns found in **F** tagged abstracts demonstrate that women are underrepresented and under-examined in the news.

7 CONCLUSION

In this paper, we have investigated that gender bias in media appears in different forms such as ideological bias, coverage bias, selection bias, and presentation bias in the news. We discussed that to secure users' attention, news titles and abstracts are typically written with contentious sentences or clauses. We conducted a pioneering initial study of implicit and explicit gender bias in news abstracts from two benchmark news recommendations and news classifications datasets, and conclude that gender bias has been present in the news and has been around for decades. By systematically conducting large scale analyses of each news corpora we detected and examined gender biases in form of (1) bias in gender distribution across all news categories and exploring the top four intersecting career words (prefixes) for *females* compared to their respective *male* counterparts; (2) bias in content in terms of attribute words which consist of 2 word categories (a) Possessive words dataset which contains a total of 465 masculine and feminine gender-specific and gender-neutral possessive nouns, and (b) Attribute words dataset which contains a total of 357 masculine, feminine and neutral career-related and family-related words; and (3) bias in wording by constructing CRA networks for the top 20 most central nouns for both gender-tagged abstracts. Each graph illustrated the compound nouns that contributed the most tagged abstracts.

[8] https://networkx.org/

Although we acknowledge that women account for half of the world's population they are incredibly under-examined and underrepresented in the news. We can immediately deduce that in both datasets, categories such as *Politics* and *Business* contain the largest measure of gender bias, as *females* are immensely under-examined and underrepresented in these areas. Male dominance is prevalent and thoroughly documented while women are depicted as '*family oriented*', and consequently we observe that news media heavily influences gender roles in society. As *females* (**F** tagged abstracts) are undoubtedly heavily associated with family words in comparison to males (**M** tagged abstracts) who are often associated with political and occupational terms. Many disciplines such as sociology, social psychology, sociolinguistics and so on – study the phenomena of how language (and written text) plays a crucial role in upholding social hierarchies.

In addition, we construct the two large benchmark datasets as follows: a possessive (gender-specific and gender-neutral) nouns and a attribute (career-related and family-related) words dataset to study the paradox of gender bias, and we will release our gendered-word datasets to foster both bias and fairness research in multiple domains such as branches in computer science and computational social science which would help to build fair NLP models by eliminating the gender bias. Although, since we focused on the **M**/**F** tagged abstracts of both news datasets, there is still a need to address the socially-constructed gender biases in the news in regards to public affairs and politics. Future works may include building fair NLP models that are trained on our two large benchmark possessive nouns and attribute words datasets. These new datasets will be monitored and updated, and therefore can be directly applied to NLP tasks such as text classification, word embeddings, coreference resolution, language modeling, machine translation, semantic role labeling, dialogue generation, etc.

REFERENCES

[1] Mohadeseh Amini and Parviz Birjandi. 2012. Gender Bias in the Iranian High School EFL Textbooks. *English Language Teaching* 5, 2 (2012), 134–147.

[2] Tolga Bolukbasi, Kai-Wei Chang, James Y Zou, Venkatesh Saligrama, and Adam T Kalai. 2016. Man is to Computer Programmer as Woman is to Homemaker? Debiasing Word Embeddings. In *Advances in Neural Information Processing Systems 29*, D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, and R. Garnett (Eds.). Curran Associates, Inc., 4349–4357. http://papers.nips.cc/paper/6228-man-is-to-computer-programmer-as-woman-is-to-homemaker-debiasing-word-embeddings.pdf

[3] Shikha Bordia and Samuel R. Bowman. 2019. Identifying and Reducing Gender Bias in Word-Level Language Models. *CoRR* abs/1904.03035 (2019). arXiv:1904.03035 http://arxiv.org/abs/1904.03035

[4] Shikha Bordia and Samuel R. Bowman. 2019. Identifying and Reducing Gender Bias in Word-Level Language Models. arXiv:1904.03035 [cs.CL]

[5] Daniel Borkan, Lucas Dixon, Jeffrey Sorensen, Nithum Thain, and Lucy Vasserman. 2019. Nuanced metrics for measuring unintended bias with real data for text classification. In *Companion Proceedings of The 2019 World Wide Web Conference*. 491–500.

[6] Aylin Caliskan, Joanna J Bryson, and Arvind Narayanan. 2017. Semantics derived automatically from language corpora contain human-like biases. 356 (14 April 2017), 183–186. https://doi.org/10.1126/science.aal4230

[7] Won Ik Cho, Ji Won Kim, Seok Min Kim, and Nam Soo Kim. 2019. On Measuring Gender Bias in Translation of Gender-neutral Pronouns. arXiv:1905.11684 [cs.CL]

[8] Steven R. Corman, Timothy Kuhn, Robert D. Mcphee, and Kevin J. Dooley. [n.d.]. Studying Complex Discursive Systems. *Human Communication Research* ([n. d.]).

[9] Jamell Dacon. 2020. Recommender System Datasets. https://github.com/daconjam/Recommender-System-Datasets.

[10] Emily Dinan, Angela Fan, Adina Williams, Jack Urbanek, Douwe Kiela, and Jason Weston. 2020. Queens are Powerful too: Mitigating Gender Bias in Dialogue Generation. arXiv:1911.03842 [cs.CL]

[11] Jaschar Domann, Jens Meiners, Lea Helmers, and A. Lommatzsch. 2016. Real-time News Recommendations using Apache Spark. In *CLEF*.

[12] Lisa Fan, Marshall White, Eva Sharma, Ruisi Su, Prafulla Kumar Choubey, Ruihong Huang, and Lu Wang. 2019. In Plain Sight: Media Bias Through the Lens of Factual Reporting. In *EMNLP-IJCNLP*. Hong Kong, China, 6343–6349. https://doi.org/10.18653/v1/D19-1664

[13] Joel Escudé Font and Marta R Costa-Jussa. 2019. Equalizing gender biases in neural machine translation with word embeddings techniques. *arXiv preprint arXiv:1901.03116* (2019).

[14] Danielle Gaucher, Justin Friesen, and Aaron C Kay. 2011. Evidence that gendered wording in job advertisements exists and sustains gender inequality. *Journal of personality and social psychology* 101, 1 (2011), 109.

[15] Hila Gonen and Yoav Goldberg. 2019. Lipstick on a pig: Debiasing methods cover up systematic gender biases in word embeddings but do not remove them. *arXiv preprint arXiv:1903.03862* (2019).

[16] Felix Hamborg, Karsten Donnay, and Bela Gipp. 2018. Automated identification of media bias in news articles: an interdisciplinary literature review. *International Journal on Digital Libraries* (2018), 1–25.

[17] C. Hutto and E. Gilbert. 2014. VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text. In *ICWSM*.

[18] Aylin Caliskan Islam, Joanna J. Bryson, and Arvind Narayanan. 2016. Semantics derived automatically from language corpora necessarily contain human biases. *CoRR* abs/1608.07187 (2016). arXiv:1608.07187 http://arxiv.org/abs/1608.07187

[19] Sen Jia, Thomas Lansdall-Welfare, and Nello Cristianini. 2015. Measuring Gender Bias in News Images. In *Proceedings of the 24th International Conference on World Wide Web (WWW '15 Companion)*. Association for Computing Machinery.

[20] Sen Jia, Thomas Lansdall-Welfare, Saatviga Sudhahar, Cynthia Carter, and Nello Cristianini. 2016. Women Are Seen More than Heard in Online Newspapers. *PloS one* 11 (02 2016), e0148434. https://doi.org/10.1371/journal.pone.0148434

[21] Andrienne Lafrance. 2016. I Analyzed a Year of My Reporting for Gender Bias (Again). https://www.theatlantic.com/technology/archive/2016/02/gender-diversity-journalism/463023/.

[22] Xiaolan Lei. 2006. Sexism in language. *Journal of Language and Linguistics* 5, 1 (2006), 87–94.

[23] Haochen Liu, Jamell Dacon, Wenqi Fan, Hui Liu, Zitao Liu, and Jiliang Tang. 2020. Does Gender Matter? Towards Fairness in Dialogue Systems. arXiv:1910.10486 [cs.CL]

[24] Haochen Liu, Wentao Wang, Yiqi Wang, Hui Liu, Zitao Liu, and Jiliang Tang. 2020. Mitigating Gender Bias for Neural Dialogue Generation with Adversarial Learning. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Association for Computational Linguistics, Online, 893–903.

[25] Kaiji Lu, Piotr Mardziel, Fangjing Wu, Preetam Amancharla, and Anupam Datta. 2020. Gender bias in neural natural language processing. In *Logic, Language, and Security*. Springer, 189–202.

[26] Juan M Madera, Michelle R Hebl, and Randi C Martin. 2009. Gender and letters of recommendation for academia: agentic and communal differences. *Journal of Applied Psychology* 94, 6 (2009), 1591.

[27] Chandler May, Alex Wang, Shikha Bordia, Samuel R. Bowman, and Rachel Rudinger. 2019. On Measuring Social Biases in Sentence Encoders. *CoRR* abs/1903.10561 (2019). arXiv:1903.10561 http://arxiv.org/abs/1903.10561

[28] Paola Medel and Vahab Pournaghshband. 2017. Eliminating gender bias in computer science education materials. In *Proceedings of the 2017 ACM SIGCSE technical symposium on computer science education*. 411–416.

[29] Ninareh Mehrabi, Fred Morstatter, Nripsuta Saxena, Kristina Lerman, and Aram Galstyan. 2019. A survey on bias and fairness in machine learning. *arXiv preprint arXiv:1908.09635* (2019).

[30] Michela Menegatti and Monica Rubini. 2017. Gender bias and sexism in language. In *Oxford Research Encyclopedia of Communication*.

[31] Rishabh Misra. 2018. *News Category Dataset*. https://doi.org/10.13140/RG.2.2.20331.18729

[32] Ji Ho Park, Jamin Shin, and Pascale Fung. 2018. Reducing gender bias in abusive language detection. *arXiv preprint arXiv:1808.07231* (2018).

[33] Alexandra Guedes Pinto, Henrique Lopes Cardoso, Isabel Margarida Duarte, Catarina Vaz Warrot, and Rui Sousa-Silva. 2020. Biased Language Detection in Court Decisions. In *International Conference on Intelligent Data Engineering and Automated Learning*. Springer, 402–410.

[34] Jan Riebling. 2015. Centering Resonance Analysis Using NLTK and NetworkX. http://www.sociology-hacks.org/?p=151.

[35] Rachel Rudinger, Jason Naradowsky, Brian Leonard, and Benjamin Van Durme. 2018. Gender bias in coreference resolution. *arXiv preprint arXiv:1804.09301* (2018).

[36] David Manning White. 1950. The "Gate Keeper": A Case Study in the Selection of News. *Journalism Quarterly* (1950).

[37] Alden Williams. 1975. Unbiased Study of Television News Bias. *Journal of Communication* (1975).

[38] Fangzhao Wu, Ying Qiao, Jiun-Hung Chen, Chuhan Wu, Tao Qi, Jianxun Lian, Danyang Liu, Xing Xie, Jianfeng Gao, Winnie Wu, and Ming Zhou. 2020. MIND: A Large-scale Dataset for News Recommendation. In *ACL*.

[39] Guanhua Zhang, Bing Bai, Junqi Zhang, Kun Bai, Conghui Zhu, and Tiejun Zhao. 2020. Demographics Should Not Be the Reason of Toxicity: Mitigating Discrimination in Text Classifications with Instance Weighting. *arXiv preprint arXiv:2004.14088* (2020).

[40] Xiang Zhang, Junbo Jake Zhao, and Yann LeCun. 2015. Character-level Convolutional Networks for Text Classification. In *NIPS*.

[41] Jieyu Zhao, Tianlu Wang, Mark Yatskar, Vicente Ordonez, and Kai-Wei Chang. 2017. Men also like shopping: Reducing gender bias amplification using corpus-level constraints. *arXiv preprint arXiv:1707.09457* (2017).

[42] Jieyu Zhao, Tianlu Wang, Mark Yatskar, Vicente Ordonez, and Kai-Wei Chang. 2018. Gender Bias in Coreference Resolution: Evaluation and Debiasing Methods. arXiv:1804.06876 [cs.CL]

[43] Jieyu Zhao, Yichao Zhou, Zeyu Li, Wei Wang, and Kai-Wei Chang. 2018. Learning gender-neutral word embeddings. *arXiv preprint arXiv:1809.01496* (2018).

[44] Pei Zhou, Weijia Shi, Jieyu Zhao, Kuan-Hao Huang, Muhao Chen, Ryan Cotterell, and Kai-Wei Chang. 2019. Examining Gender Bias in Languages with Grammatical Gender. arXiv:1909.02224 [cs.CL]

A APPENDIX

As previously mentioned, we provide one of the largest non offensive, non repeating set of gender-specific (*male* and *female*) words, we will now detail the 2 categories containing a total of 465 masculine and feminine gender possessive nouns. Note that in the creation of the set of words, overly offensive gender related words such as *bitch, whore, slut, bastard, prick, etc.,* were left out of the sets of nouns as they are hardly ever used in news articles. However, offensive gender related words are often used in tabloids (a compact version of a newspapers dominated by headline titles and images). [19].

A.1 Male Possessive Words

The succeeding word list consists of 230 gender specific words that entail *male* possessive nouns as follows:

god, gods, nephew, nephews, baron, father, fathers dukes, dad, beau, beaus, daddies, policeman, policemen, grandfather, landlord, landlords, monk, monks, step-son, step-sons, milkmen, chairmen, chairman, steward, men, masseurs, son-in-law, priest, king, governor, waiter, daddy, steward, emperor, son, proprietor, groom, grooms, gentleman, gentlemen, sir, wizards, sorcerer, lad, milk-man, grandson, grand-son, congressmen, dads, manager, prince, stepfathers, boyfriend, shepherd, shepherds, males, grandfathers, grand-fathers, husband, usher, postman, stags, husbands, host, boy, waiter, bachelor, bachelors, businessmen, duke, sirs, papas, heir, uncle, princes, fiance, mr, lords, father-in-law, actor, actors, postmaster, headmaster, heroes, businessman, boars, wizard, sons-in-law, fiances, uncles, hunter, lads, masters, brother, hosts, poet, hero, grandpa, grandpas, manservant, heirs, male, tutors, millionaire, congressman, sire, sires, widower, grandsons, grandsons, boys, he, step-father, jew, bridegroom, bridegrooms stepfather, widowers, abbot, mr., brothers, man, sons, boyfriends, he's, his, him, earl, giant, count, stepson, stepsons, poet, mayor, peer, negro, abbot, traitor, benefactor, instructor, conductor, founder, founders, hunters, huntresses, temptress, enchanter, enchanters, songster, songsters, murderer, murderers, patron, patrons, author, czar, guy, spokesman, spokesmen, pa, councilman, council-man, councilmen, council-men, gay, gays, prostate cancer, fraternity, fraternities, salesman, dude, dudes, paternal, brotherhood, statesman, statesmen, countryman, countrymen, suitor, macho, papa, strongman, strongmen, boyhood, manhood, masculine, macho, horsemen, brethren, chap, chaps, schoolboy, schoolboys, bloke, blokes, patriarch, patriachy, fatherhood, hubby, hubbies, fella, fellas, handyman, fraternal, bro, masculinity, ballerino, *pappy, papi, pappies, dada, bf, bfs, knights, knight, menfolk, brotherly, manly, pimp, pimps, homeboy, homeboys, grandnephew, grandnephew, grand-nephew, grand-nephews, john doe, nobleman, noblemen, dream boy, himself, gramps*

A.2 Female Possessive Words

The succeeding word list consists of 235 gender specific words that entail *female* possessive nouns as follows:

goddesses, niece, baroness, mother, duchesses, mom, belle, belles, mummies, policewoman, grandmother, landlady, landladies, nuns, stepdaughter, milkmaids, chairwomen, stewardesses, women, masseuses, daughter-in-law, priestesses, stewardess, empress, daughter, queens, proprietress, brides, lady, queen, matron, waitresses, mummy, empresses, madam, witches, sorceress, lass, milkmaid, granddaughter, grand-daughter, congresswomen, moms, manageress, princess, stepmothers, stepdaughters, girlfriend, shepherdess, females, grand-mothers, grandmothers, step-daughter, nieces, priestess, wife, mother, usherette, postwoman, hind, wives, murderess, hostess, girl, waitress, spinster, shepherdess, businesswomen, duchess, madams, mamas, nun, heiress, aunt, princesses, fiancee, mrs, ladies, mother-in-law, actress, actresses, postmistress, headmistress, heroines, bride, businesswoman, baronesses, sows, witch, daughters-in-law, aunts, huntress, lasses, mistress, mistresses, sister, hostesses, poetess, masseuse, heroine, goddess, grandma, grandmas, maidservant, heiresses, patroness, female, governesses, millionairess, congresswoman, dam, widow, granddaughters, grand-daughters, headmistresses, girls, she, policewomen, stepmother, stepmother, widows, abbess, mrs., chairwoman, sisters, mama, woman, daughters, girlfriends, she's, her, maid, countess, giantess, poetess, jewess, mayoress, peeress, negress, abbess, traitress, benefactress, instructress, conductress, founder, huntress, temptress, enchantress, songstress, murderess, murderesses, patronesses, authoress, czarina, spokeswoman, spokeswomen, ma, councilwoman, council-woman, councilwomen, council-women, mum, lesbian, lesbians, breast, breasts, maiden, maidens, sorority, sororities, saleswoman, dudette, maternal, feminist, feminists, sisterhood, housewife, housewives, stateswoman, stateswomen, countrywoman, countrywomen, chick, chicks, mommy, strongwoman, strongwomen, babe, babes, diva, divas, feminine, feminism, gal, gals, sistren, schoolgirl, schoolgirls, matriarch, matriarchy, motherhood, wifey, sis, femininity, ballerina, ballerinas, granny, grannies, mami, momma, maam, gf, gfs, damsel, damsels, vixen, vixens, nan, nanny, nannies, auntie, womenfolk, sisterly, motherly, homegirl, homegirls, grand-niece, grand-nieces, grandniece, grandnieces, jane doe, noblewoman, noblewomen, dream girl, madame, herself, hers

B APPENDIX

As previously mentioned, we provide one of the largest gender-specific and gender-neutral words containing a total of 357 masculine, feminine and neutral career-related and family-related words, we will now we will now detail the 2 categories of family-related and career-related words.

B.1 Career Words

The succeeding word list consists of 162 gender specific and gender neutral career-related words as follows:

policewoman, milkmaids, chairwomen, stewardesses, masseuses, priestesses, stewardess, proprietress, waitresses, congresswomen,moms, manageress, shepherdess, priestess, usherette, postwoman, hostess, waitress, spinster, shepherdess, businesswomen, actress, actresses, postmistress, headmistress, huntress, mistress, mistresses, sister, hostesses, masseuse,maidservant, heiresses, patroness, governesses, congresswoman, headmistresses, policewomen, chairwoman, maid, mayoress, peeress, traitress, benefactress, instructress, conductress, huntress, temptress, enchantress, songstress, spokeswoman, spokeswomen, councilwoman, council-woman, councilwomen, council-women, saleswoman, stateswoman, stateswomen, policeman, policemen, landlord, landlords, chairmen, chairman, steward, priest, king, governor, waiter, steward, proprietor, sorcerer, congressmen, dads, manager, waiter, actor, actors, postmaster, headmaster, businessman, manservant, tutors, congressman, benefactor, instructor, conductor, founder, founders, hunters, huntresses, tempt, enchanter, enchanters, spokesman, spokesmen, councilman, council-man, councilmen, council-men, salesman, handyman, knights, knight, academic, accountant, administrator, advisor, appraiser, architect, baker, bartender, business, career, carpenter, chemist, clerk, company, corporation, counselor, educator, electrician, engineer, examiner, executive, hairdresser, hygienist, industry, inspector, instructor, investigator, janitor, lawyer, librarian, machinist, management, mechanic, nurse, nutritionist, occupation, officer, paralegal, paramedic, pathologist, pharmacist, physician, plumber, practitioner, programmer, psychologist, receptionist, salary, salesperson, scientist, specialist, supervisor, surgeon, technician, therapist, veterinarian, worker

B.2 Family Words

The succeeding word list consists of 195 gender specific and gender neutral family-related words as follows:

niece, mother, mom, mummies, grandmother, nuns, stepdaughter, women, daughter-in-law, daughter, queens, brides, mummy, empresses, madam, granddaughter, grand-daughter, moms, stepmothers, stepdaughters, girlfriend, grand-mothers, grandmothers, step-daughter, nieces, wife, mothers, wives, girl, madams, mamas, aunt, fiancee, mrs, mother-in-law, bride, daughters-in-law, aunts, heir, heiress, sister, grandma, grandmas, dam, widow, granddaughters, granddaughters, girls, she, step-mother, stepmother, mrs., sisters, mama, woman, daughters, girlfriends, ma, mum, mommy, gal, gals, sistren, matriarch, matriarchy, motherhood, wifey, sis, granny, grannies, mami, momma, ma'am, gf, gfs, damsel, damsels, vixen, vixens, nanny, nannies,auntie, womenfolk, sisterly, motherly, homegirl, homegirls, grandneice, grand-neices, grandneice, grandneices, madame, him, father, fathers, dad, beau, beaus, daddies, grandfather, step-son, step-sons, men, son-in-law, daddy, son, groom, grooms, sir, grandson, grand-son, dads, prince, stepfathers,boyfriend, grandfathers, grand-fathers, husband, husbands, boy, bachelor, bachelors, sirs, papas, uncle, princes, fiance, mr, father-in-law, sons-in-law, fiances, uncles, brother, grandpa, grandpas, widower, grandsons, grand-sons,boys, step-father, bridegroom, bridegrooms, stepfather, widowers, mr., brothers, man, sons, boyfriends, he's, his, stepson, stepsons, guy, fraternity, fraternities, salesman, dude, dudes, paternal, brotherhood, papa, boyhood, manhood, masculine, brethren, chap, chaps, patriarch, patriachy, fatherhood, hubby, hubbies, fella, fellas, fraternal, bro, pappy, papi, pappies, dada, bf, bfs, brotherly, homeboy, homeboys,grandnephew, grandnephew, grand-nephew, grand-nephews, gramps, family, infancy, infant, kin, orphan, twin