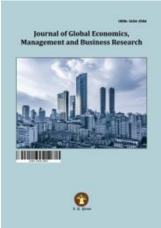


Journal of Global Economics, Management and Business Research

Volume 16, Issue 1, Page 19-38, 2024; Article no.JGEMBR.12023
ISSN: 2454-2504

Unveiling Gender Biases in Recruitment: A Natural Language Processing Approach

Mirian Izquierdo Barriuso [a*]

[a] *University of Witten-Herdecke, Germany.*

Author's contribution

The sole author designed, analysed, interpreted and prepared the manuscript.

Article Information

DOI: 10.56557/JGEMBR/2024/v16i18634

Open Peer Review History:
This journal follows the Advanced Open Peer Review policy. Identity of the Reviewers, Editor(s) and additional Reviewers, peer review comments, different versions of the manuscript, comments of the editors, etc are available here: https://prh.ikprress.org/review-history/12023

Review Article

Received: 26/01/2024
Accepted: 30/03/2024
Published: 08/04/2024

ABSTRACT

This paper investigates the potential of AI to identify gender biases in recruitment for senior management positions in businesses, dealing with many documents. It aims to unravel the impact of gender biases in job advertisements as a possible reason behind the underrepresentation of women in the corporate world. An innovative experiment that extracts and analyses 2.198 job offers published in February and September 2021 in the Financial Times newspaper is presented. Natural language techniques are used. These methods identify the most frequent terms and their appearance rate in the advertisements showing the gender biases they generate. By enabling the analysis of many documents, the method allows the accurate identification of gender biases. This use is unique in management studies. The results show a strong co-occurrence of terms associated with male roles in the studied sectors. The concept of agentic-communal role differentiation, rooted in the Identity, homosocial, and TM-TM theories, supports the findings. This knowledge will contribute to using NLP to discover gender biases in recruitment for high decision-making positions proposing action to improve the present situation. Should more women ascend to decision-making positions, selection processes should be improved to reflect a more neutral language. At the same time, cultural changes should be promoted in the corporate world toward more inclusive workplaces.

**Corresponding author: E-mail: presidencia@womanforward.org;*

Keywords: Natural language processing techniques; social identity; homosocial and TM-TM theories; job postings; recruitment.

1. INTRODUCTION

In recent decades, there has been a cultural change in the corporate world regarding gender equality, characterized by the significant increase in the number of women who have received professional training to enter and develop their careers. This cultural change has led society to demand businesses greater social responsibility and, consequently, changes in the legal framework through laws and measures aimed at achieving real and effective equality between women and men.

Despite the progress that has been made, there are still areas for improvement to guarantee effective equal opportunities. The fact that women face more obstacles in reaching decision-making positions in the case of selection and recruitment processes, is due to multiple factors. Most have to do with sexist prejudices related to gender stereotypes and biases leading to discrimination [1,2].

All considered, gender biases in recruitment remain an issue for society and the economy. Social fairness and equity are harmed, but innovation, diversity, and economic growth also suffer. By recognizing and addressing the negative influence of these biases, businesses can cultivate inclusive environments that will benefit both society and the individual.

Job postings are the first step for recruitment and, thus, the door to these potential biases. Job posting for high decision-making positions can vary depending on the industry and the specific job role. However, there are some elements that are common. Specifically, stereotypes persevere in the frequent gendered language, trait emphasis, and biased requirements of job postings.

AI can play a decisive role in identifying gender biases by dealing with a great source of documents. NLP techniques can analyse terms and sets of words to identify the language and traits associated with managerial positions and then compare how those traits are associated with gender. These techniques place themselves, therefore, as important and innovative in the path to improve diversity in recruitment.

As the requirements are listed, male candidates may find more advantages in certain qualifications. Aspects unrelated to performance, such as high emphasis on leadership style, previous job title, or years of experience, benefit male applicants more [3]. The job title may include words like "Director," "Vice President," "Manager," or "Executive," depending on the level of the position. It will likely list the key responsibilities and duties of the position, which include making critical decisions that impact the organization's strategy and direction, developing and implementing plans to achieve specific goals, managing budgets and resources, and leading teams.

In the context of this research, Social Identity Theory, Homosocial, and the TM-TM theories can be applied to explore how job postings´ wording may influence recruitment processes for high decision-making positions.

This paper aims to unravel the impact of gender biases in job advertisements as a possible reason behind the underrepresentation of women in the corporate world. The main objective is, thus, if the effectiveness of using NLP techniques adds a new dimension to identifying gender biases from job postings.

The rest of the paper is sorted into four different sections. In the first section, we present the theoretical background that supports the development of the hypothesis. I explain the main concepts and theories that may apply to identify biases in job advertisements.

The methods and analyses follow right after in the third section. In the analyses, I use vectorization techniques to reduce documents to the vector of words; frequency technique, a metric in the field of information retrieval (IR), which weights the frequency of occurrence of each term found considering the whole set of documents of the job postings, and the Silhouette method to make homogenous clusters of documents. We then use the Semantic distance of each cluster to words indicating male or female roles, following a traditional agentic versus the communal list of terms.

In the findings and discussion section, I summarize key findings and interpret the results. I then discuss the implications, strengths, and limitations of the study. In the conclusion section, I explain how NLP can identify gender biases and contribute to gender equality in organizations.

2. CONCEPTS, THEORIES, AND HYPOTHESES

In this paper, I study gender biases in job offerings by applying NLP techniques. This section clarifies concepts and theories that may contribute to discussing the effectiveness of NLP techniques and the extent to which they can contribute to identifying gender biases in job postings.

I then present the impact of gender biases in job postings in high decision-making positions and their impact on diversity in recruitment methods. The concepts of Communal and Agentic Roles and Values are presented to understand how the use of gendered language may reduce and may have an impact, on enduring gender segregation in the workplace. Hypotheses are then presented, and theories are to explain how and to what extent gendered language impacts selection processes.

2.1 Natural Language Processing, Text Classification, Soft Computing Techniques: Its Use and Limitations in Recruitment

Natural Language Processing (NLP) can be defined as a subfield of artificial intelligence that focuses on the interaction between natural human language and computer systems. It involves developing algorithms and models that can analyse, understand, and generate human language in a way that is useful to detect gender biases and process a large quantity of data.

The ability to screen large quantities of information is a strength of the automated DS systems that can make selection processes more time-efficient [4]. There is, therefore, an increasing interest in using automated decision support in personnel selection [5].

In addition to that, HR managers could benefit from automated tools in selection without the need to be exposed to negative perceptions by important stakeholders [6].

Various NLP techniques and algorithms, such as machine learning algorithms and lexical analysis, can be used to analyse and generate natural languages [7]. If they are robust, generalizable, and can handle variation and ambiguity, they may be used to analyse the language used in job postings. Providers of AI-based personnel selection solutions commonly market their systems in a way that highlights the potential for less bias in personnel selection when organizations use their systems [8].

However, several authors maintain that more evidence that such systems can prevent bias is still needed. According to research, it might be necessary to train decision-makers on challenges when relying on automated systems [9]. This training becomes crucial if human decision-makers are responsible for overseeing AI-based systems in the selection, as proposed in current drafts for legislation on AI (e.g., the European AI Act). A working definition of biased systems can be explained as “computer systems that systematically and unfairly discriminate against certain individuals or groups of individuals in favour of others” [10]. Additionally, it has been pointed out that the modes of algorithmic knowledge production through categorization and classification are needed to better understand their workings and the resulting biases [11].

Bohnet presented that behavioural design could be used to reduce gender biases in job postings [12]. She suggested that AI and NLP could replace gendered wording, improving more neutral job posting advertisements. Additionally, a methodology was developed to remove gender stereotypes from word embeddings [13], a natural language processing technique that reduces the mathematical dimensionality of words or phrases’ representations and thus makes them easier to process computationally. Fairness criteria like demographic parity [14] and equal opportunity [15] prohibit discrimination against protected attributes on a statistical level in the internal model of an algorithm, avoiding, thus, redundantly encoding. A study in 2018 provided a demonstration that AI algorithms can identify biases in texts using word embedding to analyse gender and ethnic biases in different types of texts, including job postings [16]. Garg showed that biases had decreased substantially, but some persist.

Recently, researchers sustained that soft computing techniques, such as fuzzy logic and neural networks, can be used to analyse and classify natural language data and, thus, to detect biases in automated systems. Like NLP, soft computing techniques and text classification can also be employed to boost equity and identify biases in automated systems due to their ability to organize and analyse natural language data. If these systems were to apply bias-aware algorithms to examine training data for biases and attempt to lessen them, they could be executed to avert system bias [17].

By using NLP techniques to analyse the language used in job postings, it may be possible to identify patterns of gendered language. Moreover, by implementing a bias-aware learning algorithm that explicitly considers the potential biases in the training data and attempts to mitigate them, it can be testes whether these patterns have a positive impact on the gender composition of job applicants. I, therefore, hypothesize as follows:

2.2.1 Hypothesis 1

NLP techniques in recruitment processes for high decision-making positions are helpful in identifying gender biases, and potentially increase a more diverse pool of candidates.

2.2 Gender Biases in Job Postings and Theories

The literature on gender has traditionally identified gender biases in job postings and their consequences for high decision-making positions. I present here the theories that further explain the consequences of these biases on women´s careers.

Gender differences in the linguistic style of every day have been extensively researched [18,19]. Women use a more communal style of speech than men [20,21,22] and make more references to social and emotional words [23].

Despite progress toward gender equality at work in recent years, gender segregation and clear gender differences in labour market outcomes persist [24]. The hiring process is critical in addressing gender inequality [25]. Ibarra and others explore the impact of gender-exclusive language in job postings on women's feelings of belonging and interest in applying for high decision-making positions [26,27]. Stout and Dasgupta posit that gender biases in job postings for high decision-making positions include the underrepresentation of women in leadership roles, perpetuating gender stereotypes, self-selection out of opportunities, and gender inequality in the workplace [28].

Research has provided clear evidence that job postings, explicitly and implicitly, signal employers' preferences for certain types of candidates, which may lead to potential gender biases. There is an indication that the language and wording used in UK STEM job advertisements are biased toward a masculine orientation [29]; and even that there is a positive association between male-biased job postings and a male-dominated STEM workforce [30].

2.3 Communal and Agentic Roles and Values

Understanding agentic and communal roles is crucial for recognizing and addressing gender stereotypes in decision-making positions and how job posts may be impacted. By acknowledging these roles and their influence on our perceptions, work can be done towards reducing gender biases. Acting 'like men' or acting 'like women' have been stereotypically introduced as agentic (male-type) and communal (female-type) behaviour.

Eagly's Social Role Theory is particularly useful when aiming to understand gender bias in job postings [31]. Through the lens of social role theory, societal rules, historical circumstances, and cultural patterns can be studied to determine the impact of language and job requirements in career posting. In this way, it can be understood whether certain job postings include language and requirements that cater to one sex due to gender stereotypes. Biases may have consequences of deterring applicants of the opposite sex, therefore creating more gender bias.

Useful for discovering gender values in job postings, Cejka and Eagly noted that gender roles are mainly created because of the difference in labour associated with men and women [32]. These different roles are connected to distinct characteristics. Agentic characteristics (traditionally associated with men) include competitiveness, assertiveness, and independence. In contrast, communal characteristics (traditionally associated with women) include cooperation, warmth, and nurture. This division encourages stereotypes as society associates agentic behaviours with men and communal behaviours with women [33]. More recently, the concept of agentic and communal designations in values and roles has gained recognition and has been employed in multiple areas of study, such as career development [34, 35]. The list of words and values of communal versus agentic roles and values is fluid. The evolution of this list represents the progressing comprehension of gender stereotypes and their influence on women´s career development [36]. This is particularly true in occupations that are typically male dominated, as female applicants are less likely to apply due to the gendered language [37].

In relation is the idea of the anticipatory sorting process. For example, it has been discovered that when women identify gender bias in a job posting, they may self-select themselves out of these positions. As happens often when women examine high decision-making positions. This idea is expanded on by Fernandez-Mateo and King [38], who conclude that this self-selection disqualification is an essential contributor to the lower levels of women in leadership roles and enduring gender segregation in the workplace.

Similarly, Bohnet, argues that not only does gender biases in the recruitment process depress the number of women applying to high decision-making positions, but it also contributes to gender stereotypes, gender inequality, and a lack of diversity [12]. When confronted with gender biases and exclusive language in a job posting, women interpret that they are not valued in these positions [28]. Therefore, they are less likely to apply, and consequently, there is a low percentage of female composition in leadership or executive positions. As can be predicted, research shows that men are attracted to roles with more agentic language, and women are attracted to roles with more communal language [39].

2.4 Homosocial Role Theory

The theory that people of the same gender form bonds with others who are related to them regarding gender is known as homosocial role theory [40]. This theory provides a significant foundation for exploring gender bias in recruitment. A result of this theory in the practical world is the skewed hiring and promotion procedures and the continuance of agentic and communal gender vocabulary in job postings. Consequentially, the male-dominated high decision-making positions are more likely to be imbalanced, as women are isolated from these places by homosociality. The extensive effects of these social bonds can be seen in workplace power structures through gendered relationships and networking, with gender-exclusive social networks and unofficial preferences for hiring. In addition to the barriers presented by homosocial role theory, the gender stereotypes of management being a masculine occupation contribute to fewer women ascending to management positions or being recruited [41]. Ollilainen & Calasanti [42] present a wider variety of values and skills in leadership, the perpetuation of a gendered hierarchy of authority, and the use of metaphors with teams. Holgersson's research [43] on the recruitment of managing directors reveals that homosociality can contribute to gender biases in job postings and hiring processes.

2.5 Think-Manager Think-Male Theory

The TM-TM association [44] has been one of the core arguments in the discussions of women´s empowerment in the business world. TM-TM is founded on a belief that effective management is based on ‘men-like’ or male-type values.

Gender stereotypes about desired management qualities are linked through cyclical reinforcement. A study in 2001 by Schein expanded on this connection, as society associates masculine traits with necessary leadership traits. This overlap of characteristics between masculinity and leadership includes qualities like assertiveness, dominance, and decisiveness. Therefore, leadership positions emphasize these traits, which men identify with. The high percentage of men in these positions continues to strengthen the stereotype that these positions are more suited for men, and future posts include the same gendered language. The concept of the Think-Male-Think-Manager theory is explored more with the link between gendered traits and forms of leadership [45].

Managers are recruited and rewarded based on their capacity to act goal-driven, decision-making, and assertively. Oppositely, ‘women-like’ or female-type leadership values typically refers to acting oriented to others and their well-being [46,47].

The three theories of Social Theory, Think-Male-Think-Manager Theory, and Homosocial Theory are relevant material for analysing gender bias in job recruitment. They can assist thus in the outcome of a basis for communal versus agentic language. The following Fig. 1, can be used as a representation of agentic and communal roles based on different contributions.

Based on this, I hypothesize as follows:

2.5.1 Hypothesis 2

The co-occurrence of gendered terms in job postings varies within sectors. The occurrence of gender-biased language and masculine-coded terms is higher in more agentic sectors.

Male role agentic	Female role communal
Performance	*Collaborative*
Problem	*Responsibility*
Strategy	*Equality*
Decision	*Reputation*
Team	*Transparency*
Market	*Security*
Consumer/client	*Ethics*
Growth	*Disclosure*
Pressure	*Empowerment*
Power	*Sponsorship*
Outcome	*Social*
Result	*Justice*
Financial	*Sensitivity*

Fig. 1. Agentic and communal word comparison

3. METHODS AND FINDINGS: EXTRACTION, PRE-PROCESSING, AND ANALYSIS OF INFORMATION

This section provides a description of the methods used to analyze the job advertisements, specifying the role of the Bag of Words model in uncovering the linguistic trends within the dataset of the analysis. It balances technical detail with accessibility, ensuring that readers from various disciplines can grasp the methodological approach and its relevance to the study's objectives.

A structured exposition delineates the analytical techniques employed and contextualizes their application within the broader objectives of the study, ensuring that the methodology section is comprehensive and coherent. It integrates the use of word clouds as a method for qualitative analysis within the framework of the research, providing a detailed description of the findings and their implications for understanding the composition and focus of the job advertisement clusters. It furthermore integrates the analysis of gendered language within the clusters, explaining the methodology employed to quantify the association of words with gendered values. It finally provides a nuanced interpretation of the results, situating the findings within a broader socio-linguistic and theoretical context.

3.1. Data Description

The methodology of this study was predicated on the systematic collection and analysis of job advertisements published by the Financial Times on two specific dates: February 28, 2021, and September 4, 2021. A comprehensive dataset was compiled from a total of 2,198 advertisements, each meticulously catalogued within a purpose-built database. This database was structured to include the following fields for each advertisement: a unique identifier (ID), the job location, the salary range offered, the recruiter responsible for the hiring process, a URL to the full advertisement, and the advertisement's full text as accessed through the provided URL.

The principal aim of this endeavor was to identify and analyze the prevalence and significance of key terms within each job offer. Such analysis is intended to illuminate the most frequently emphasized attributes or qualifications sought by employers.

3.2 Information Processing

The initial phase of data preparation involved a rigorous cleaning process to ensure the uniformity and quality of the dataset, thus facilitating more accurate and insightful analysis. Critical to this process was the removal of stop words -words which, though frequently occurring, hold minimal analytical value (e.g., "a", "the", "and"). This was accomplished utilizing the Natural Language Toolkit (NLTK) library, version 3.5. Subsequently, a lemmatization procedure was implemented to consolidate variant forms of a word into its base or lemma form, thereby streamlining the dataset and enhancing the analytical focus on meaningful keyword frequencies.

Punctuation was eliminated to further reduce noise in the textual data. The final preprocessing step involved the tokenization and syntactical parsing of the advertisement texts, selectively retaining nouns and adjectives for subsequent analysis, resulting in a set of words as shown in Fig. 2.

3.3 Representation Models

This subsection delineates the computational strategies deployed to analyze the corpus of job advertisements, with a particular focus on evaluating the lexical features and thematic content of the advertisements. The objective is to elucidate the linguistic patterns underlying the job offers by identifying the most frequently occurring and contextually significant terms. This endeavor involves the application of several representation models to dissect the semantic structure and frequency distribution of terms within the dataset.

3.3.1 Bag of words

A foundational technique in the realm of document representation is vectorization, wherein textual data is transformed into numerical vectors that reflect the presence and contextual relevance of words within the documents. Among the various models leveraging this approach, the Bag of Words (BoW) model stands out for its simplicity and efficacy [48]. The BoW model disregards the order of words but maintains a focus on the frequency of their occurrences across the document corpus.

In applying the BoW model to the dataset, each job advertisement is reduced to a vector representing the aggregate occurrence of words within it. This process facilitates a quantitative analysis of word frequencies, enabling the identification of the most prevalent terms across all job offers. Fig. 3 illustrates the outcomes of this analysis, showcasing the words that appear most frequently along with their respective occurrence counts:

To better visualize the results, it was complemented with NLTK's FreqDist tool, as shown in Fig. 4, and with a word map or wordcloud [49], as shown in Fig. 5, which allows to better highlight the most repeated terms within the set of documents.

As seen in Fig. 4, the ten most repeated words are 'experience', 'business', 'management', team', 'role', project', 'delivery', 'financial', 'development' and 'client', with an incidence of over 600 occurrences in the total number of documents.

3.3.2 Term frequency-inverse document frequency (TF_IDF)

Following the application of the Bag of Words model, I further refined the analysis of document content through the implementation of the Term Frequency-Inverse Document Frequency (TF-IDF) technique. This metric, foundational in the field of information retrieval, evaluates the significance of a word within a document relative to a corpus, effectively balancing term frequency against its distribution across documents [50,21]. The TF-IDF score increases proportionally to the number of times a word appears in the document but is offset by the frequency of the word across the document set, thus highlighting words that are more specific to a document.

Fig. 2. Example of the set of words at the end of the text cleaning process

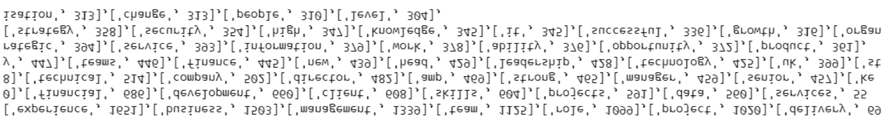


Fig. 3. Most repeated words in the total number of documents, together with their rate of occurrence

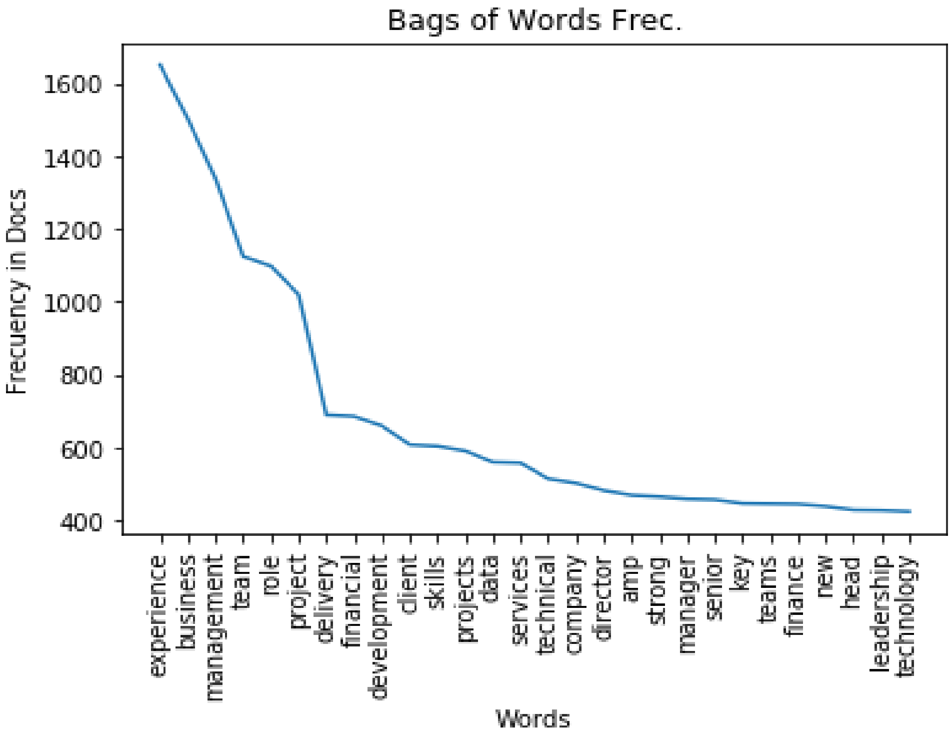


Fig. 4. Result of the bag of words distribution with the Freq Dist tool

Fig. 5. Representation model of the bag of words with the Wordcloud tool

Utilizing the NLTK library (version 3.5), I employed two algorithms to calculate TF-IDF values:

- Term Frequency (TF): This component quantifies the prevalence of a term within a specific document, calculated as the ratio of the term's occurrence to the total number of terms in the document:

TF(t) = Total number of terms in the document / Number of times term t appears in a document

- Inverse Document Frequency (IDF): This measure assesses the overall importance of a term within the entire document corpus, inversely related to the term's frequency across documents:

IDF(t) = loge (Total number of documents/Number of documents containing term t)

The amalgamation of TF and IDF values enabled the identification of terms that are not only frequent but also pivotal within individual documents. This analysis allowed for the extraction of significant terms from each cluster, facilitating a comparison with results derived from the Bag of Words approach.

3.3.3 Clustering of documents according to their most relevant terms

Capitalizing on the insights provided by TF-IDF metrics, I proceeded to cluster the job advertisements based on the relevancy and frequency of terms within them. This phase aimed to ascertain the thematic homogeneity and content-based similarity across the dataset. Employing the TF-IDF output as a foundational dataset, I applied clustering algorithms to categorize the documents into coherent groups.

The determination of an optimal cluster count was informed by the Silhouette method, which suggested a bifurcation of the dataset into two distinct clusters. This recommendation was predicated on the observed homogeneity within these clusters, which surpassed that of any alternative partitioning scheme involving a greater number of clusters. Fig. 6 illustrates the silhouette scores associated with this clustering decision, underscoring the methodological rigor and empirical basis underpinning the choice.

In this case, with two clusters, the characteristics of each document has been reduced to the 8 most relevant terms in its content, obtaining 1853 documents in the main cluster and 345 in the second cluster, as seen in Fig. 7.

To further delineate the thematic essence and terminological distinctiveness of each cluster identified within the corpus of job advertisements, I employed visual analytics, specifically through the generation of word clouds. Word clouds serve as an intuitive means to visualize the relative prominence of terms within a textual dataset, where the size of each term in the cloud is proportional to its frequency across the documents within a cluster. This methodological approach enhances the understanding of the predominant themes and keywords characterizing each cluster and aids in the qualitative analysis of the content's alignment with specific industry sectors.

Fig. 8 presents the word clouds generated for two principal clusters derived from my analysis:

- Cluster 0 (Fig. 8a): Encompassing a majority of the dataset with 1,853 documents, the word cloud for Cluster 0 reveals a broad spectrum of terms that are prevalent across the general corpus of job advertisements. The visualization highlights the diversity of skills, roles, and attributes that are commonly sought across a wide range of industries, thereby characterizing the cluster as encompassing a generalist view of the job market.

Cluster 1 (Fig. 8b): In stark contrast, Cluster 1, comprising 354 documents, is distinctly marked by terminology germane to the financial sector. The word cloud for this cluster distinctly features keywords related to finance, banking, and investment, suggesting that the documents within this cluster likely represent job offers from a homogenous sector. The specificity of the terms within this visualization indicates a concentrated grouping of offers that share a common linguistic and thematic identity, underscoring the efficacy of the clustering process in segregating documents according to their content and contextual relevance.

The utilization of word clouds as a visual exploration tool thus, substantiates the clustering analysis, providing a clear, graphical representation of the linguistic patterns that differentiate the clusters. Through this approach, it can be inferred the predominant themes and sectoral focus of the job advertisements within each cluster, thereby offering insightful perspectives into the dynamics of the job market as captured through the lens of the dataset, as Figs. 9a and 9b show.

Thus, as seen in Fig. 9a, in cluster 0 with 1853 documents, we can see the terms that characterize most of the analysed offers. However, in Fig. 9b, with 354 documents, words related to the financial sector appear, which may correspond to a group of offers belonging to the same sector that use the same more specific terms for their definition.

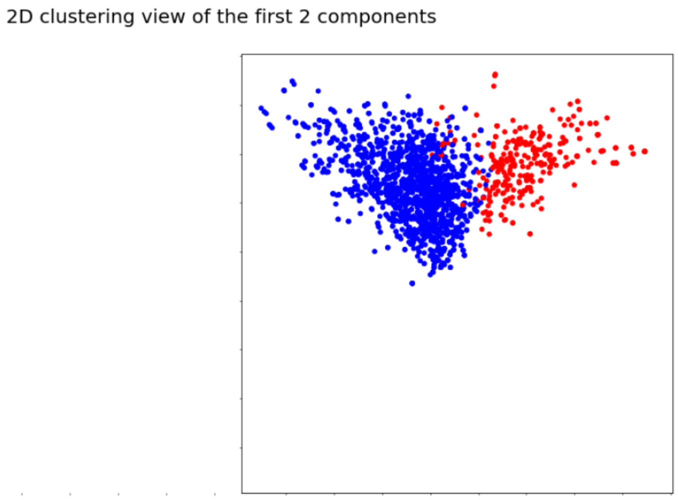
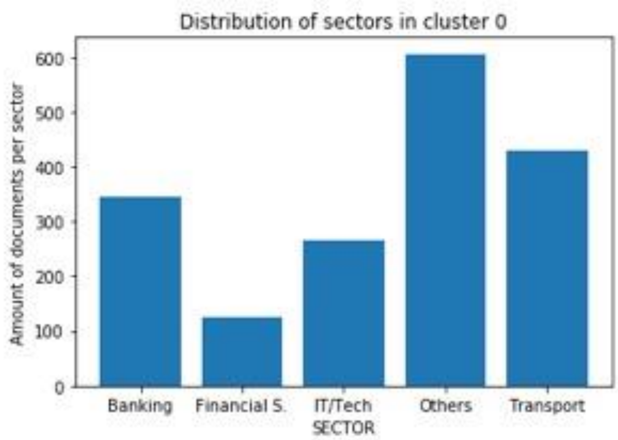


Fig. 6. Output of the silhouette method for different number of clusters

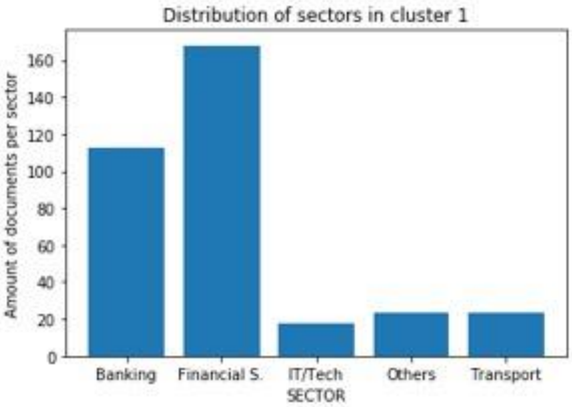
```
   Cluster         0         1         2         3         4         5  \
0        1  0.312314  0.386244  0.044064  0.087343  0.164808 -0.108293
1        1  0.281262  0.387142  0.042880  0.098653  0.145126 -0.201974
2        1  0.312314  0.386244  0.044064  0.087343  0.164808 -0.108293
3        1  0.281262  0.387142  0.042880  0.098653  0.145126 -0.201974
4        1  0.312314  0.386244  0.044064  0.087343  0.164808 -0.108293

          6         7
0 -0.099368  0.018488
1 -0.300684  0.145299
2 -0.099368  0.018488
3 -0.300684  0.145299
4 -0.099368  0.018488
         Cluster
Cluster
0           1853
1            345
```

Fig. 7. Grouping of documents according to their most relevant terms



Sector	Nº OF documents in cluster 0
IT/Tech	265
Transport	431
Financial S.	124
Banking	346
Others	606



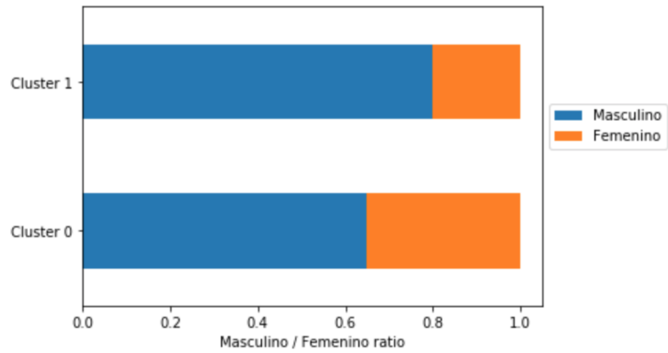
Sector	Nº OF documents in cluster 1
IT/Tech	18
Transport	23
Financial S.	168
Banking	113
Others	23

Figs. 8a and 8b. Distribution of sectors in cluster 0 (7a) and in cluster 1 (7b)

Figs. 9a and 9b. Wordclouds of the most relevant terms of the documents are contained in cluster 0 (8a) and in cluster 1 (8b)

Male role	Female role
Performance	Collaborative
Problem	Responsibility
Strategy	Equality
Decision	Reputation
Team	Transparency
Market	Security
Consumer/client	Ethics
Growth	Disclosure
Pressure	Empowerment
Power	Sponsorship
Outcome	Social
Result	Justice
Financial	Sensitivity

Fig. 10. Division of roles by gender




```
result
{'Cluster 0': {'Rol Masculino': 0.65, 'Rol Femenino': 0.35},
 'Cluster 1': {'Rol Masculino': 0.8, 'Rol Femenino': 0.2}}
```

Fig. 11. Clusters of terms co-occurrence with terms associated with communal (feminine) or agentic (masculine) values

Sector	Number of Companies	Percentage of Companies	Average Percentage of Female Board Members	Average Percentage of Female CEOs
IT/Tech	3	~ 3%	41%	3.4%
Transport	5	~ 5%	41.6%	3.25%
Financial	13	~ 14%	44.8%	6.1%
Banking	5	~ 5%	41.9%	4.9%
Others	71	~ 73%	40.7%	3.6%

Fig. 12. Results of feminization percentage of high decision-making positions by sectors in FTSE 100 companies

3.3.4 Semantic distance of each cluster to words indicating male or female roles

To assess the presence and potential bias towards gendered language within the identified document clusters, I applied a linguistic analysis grounded in the theoretical framework of communal versus agentic values. This framework, elaborated upon by scholars such as Eagly [31], Brownlow [20], Rudman & Phelan [36], Gaucher et al. (2011), and Gupta et al. (2018), distinguishes between language that conveys communal values—often stereotypically associated with femininity (e.g., supportive, collaborative) and agentic values—typically linked to masculinity (e.g., competitive, independent).

Utilizing the key terms extracted from each cluster (as visualized in Fig. 8), I calculated the semantic distance between the most relevant words within each cluster and a predefined list of terms identified as either masculine or feminine based on their association with communal or agentic values, respectively. This analysis was performed by evaluating the semantic similarity of cluster terms to each set of gender-associated words, thereby categorizing each term as 'masculine' or 'feminine' according to its closest semantic proximity to the gendered word lists.

The determination of a word's gender association was based solely on the maximum similarity score. For instance, if a given term exhibited a 0.70 similarity to a masculine-associated word and 0.20 to a feminine-associated word, it was classified under masculine due to the higher score. This methodological approach allowed for a clear, quantitative assessment of the gendered language propensity within each cluster.

My findings reveal a pronounced co-occurrence of terms aligned with masculine roles within the clusters. This indicates a potential skew towards agentic values in the language used within the job advertisements, suggesting an underlying gender bias in the portrayal of roles and expectations. Such a bias could reflect broader societal stereotypes about gender and work, as well as influence the attractiveness of job advertisements to potential applicants based on their gender identification.

Fig. 11 shows the roles in each cluster. The findings evidence that there is, in both clusters, a strong co-occurrence of terms associated with agentic values. In the case of cluster 1, the co-occurrence reaches 80%, whereas in cluster 0, the co-occurrence is 60%.

Findings show thus that hypothesis 1 is supported. NLP techniques have been helpful in analysing many documents, and identifying the gendered language used in job postings. NLP techniques are, therefore, potentially useful for improving the talent pool of candidates.

Hypothesis 2 is, to the contrary, not supported. Both clusters present a predominant use of agentic terms, although “others” and “transport”, are not specifically agentic. All sectors include gendered language when seeking candidates for high decision-making positions. Thus, the use of agentic terms, may not variate in different sectors.

In addition to that, cluster 1 presents a co-occurrence of 80% of agentic terms, representing 345 documents, that include all 5 sectors of the study. In contrast, cluster 0, presents a co-occurrence of 60% of agentic terms of 1772 documents, which include all five sectors with more documents. Based on these results, it cannot be concluded that more communal sectors, may use a more communal language.

4. DISCUSSION

Can NLP techniques be used to identify gender biases from job postings and improve recruitment diversity for high decision-making positions? What do the findings mean in practice? Through this paper, I contribute to identifying gender biases in job postings, even when dealing with extra-large samples, validating NLP techniques. I have also revisited the use of communal and agentic concepts and values in today’s advertisements for high decision-making positions.

The section communicates the specific findings of the gendered language analysis and contextualizes the findings within broader discussions about gender, work, and organizational practices, using current literature. It offers a clear, concise explanation of the results while emphasizing the implications for inclusivity and diversity in recruitment of decision-making positions.

4.1 Findings related to Gendered Language Analysis

The Findings of this research into the gendered dimensions of linguistic expression within job advertisements, yielded compelling evidence of a pronounced bias towards agentic language. The linguistic bias aligns with stereotypically masculine values, underscoring a potential gendered framing of job roles and expectations. The analysis, predicated on the measurement of semantic distances between cluster-specific terms and established lists of gender-associated words, revealed a significant skew towards agentic descriptors in both clusters examined.

Cluster 1 Analysis: In Cluster 1, which is predominantly composed of advertisements related to the financial sector, an overwhelming 80% of the terms analyzed exhibited a closer semantic affiliation with agentic, or masculine-associated, values. This finding suggests a strong inclination towards portraying roles within this cluster through a lens of leadership, competitiveness, and independence—qualities traditionally valorized in masculine narratives.

Cluster 0 Analysis: Cluster 0, characterized by a broader spectrum of job advertisements, demonstrated a 60% co-occurrence of terms aligned with agentic values. While less pronounced than in Cluster 1, this bias towards masculine language indicates a general trend across the job market, where roles are framed in terms that implicitly favor agentic over communal attributes.

4.2 Implications of Gendered Language Bias

The evident bias towards agentic language within job advertisements raises important considerations regarding the inclusivity and diversity of the workplace. The predominance of masculine-associated terms could potentially deter female candidates from applying for roles that appear to be implicitly coded as 'masculine'. This effect, known as stereotype threat, may contribute to the perpetuation of gender imbalances within certain sectors or roles.

Furthermore, the findings underscore the necessity for organizations to critically assess and possibly recalibrate the language used in job advertisements to foster a more inclusive recruitment strategy. By promoting a more balanced representation of agentic and communal values, employers can enhance the appeal of their job offers to a broader demographic, encouraging diversity and mitigating implicit gender biases.

The study has its strengths, but there are also some limitations. Limitations of this research relate to the geographical context. The vacancies analysed correspond only to the United Kingdom, as it should be kept in mind that vacancies for top decision-making positions are not normally published in other countries. This is a limitation and a finding that reflects the potential lack of objectivity in the selection of decision-making positions even today. The fact that many of the positions are not published implies a potential barrier for the least represented sex when accessing the decision-making positions under the light of the presented theories, as discussed further down.

The research contributes, therefore, to unravel the impact of gender biases in job advertisements as a possible reason behind the underrepresentation of women in the corporate world.

In the first part, I present why gender biases remain a significant issue within the corporate world, particularly in recruitment. Multiple issues develop because of these biases, effectively harming both the businesses and the women that are applying to them. When women are denied, decision-making positions based on gender biases, companies engender a detrimental atmosphere where groupthink prevails over diversity and innovation [51,52]. While creation and productivity are stifled internally, women experience the institutionalization of their supposed insufficiency. With these rejections, society continues to see women as incapable of holding certain job titles [53]. Therefore, the stereotype that women are inadequate is reproduced.

In addition to internal problems within a business, companies are beginning to experience higher social repercussions when issues like gender biases go unaddressed. Success is becoming more often awarded to organizations that participate in the acknowledgement of societal issues. In recent years, investors and customers have given more value to ethically responsible businesses [52].

For example, certain languages might imply a specific gender based on societal norms. Masculine-coded language is common in decision-making roles, such as with the words “dominant” and “assertive”. This identifies the role more with men than with women. In contrast, feminine-coded words, such as “nurturing” and “collaborative”, call more to women. However, in the case of the latter example, feminine-coded language is less common in decision-making job posts. These commonplace principles, while seeming benign, can be rooted in gender bias.

Findings reveal that words such as "pressure," "power," and "performance" are commonly used by today's corporations to describe managerial roles and that these words are more often associated with male candidates than female candidates. The strong co-occurrence of terms associated with male roles suggests that gender biases and stereotypes influence the recruitment process, potentially underrepresenting women in high decision-making positions [54]. Therefore, NLP techniques, being able to identify these words in many documents, are instrumental in addressing gendered language. Identifying this gendered language helps break the traditional perpetuation of gender inequality in the workforce. Moreover, it may change the discouragement of women from pursuing male-dominated roles and stop reinforcing the stereotype that these positions are better suited for men, explained by research [37].

Furthermore, research shows that challenges to stereotypes are necessary to advance gender equality across multiple cultures [41]. After analysing the social connection between masculine traits and development in women´s careers, it was determined that high management positions are associated with masculine characteristics in multiple countries and cultures. Thus, women face more obstacles globally when attempting to reach these managerial jobs. In addition to this barrier in the physical world, women also confront the psychological repercussions of low female representation and association in leadership. NLP techniques offer a new challenge to these stereotypes.

The gendered biases identified by NLP techniques can help organizations to work further, creating more inclusive job postings by using gender-neutral language. Schein [41] similarly suggested the need for corporations to go through cultural changes, creating more inclusive workplaces.

Similarly, Homosocial theory explains how women are kept away from high decision-making positions, resulting in gender imbalances in leadership roles [43]. The gender-exclusive language can convey to women that they are not the target of these positions, leading to their underrepresentation in decision-making positionsn [28].

If an organization's leadership is primarily male, they may be more likely to rely on their networks to fill job openings. This can lead to replicating existing gender biases [55]. Additionally, hiring managers may unconsciously favour candidates who are similar to themselves or whom they perceive as fitting in with the existing organizational culture, which can further contribute to gender biases in job postings and hiring processes. Education of hiring managers is needed to be sensitive to the need to identify gendered language.

Social role theory encompasses a large variety of viewpoints attempting to realize human behaviour through social processes and organizations, highlighting the need to go over cultural changes able to create more diverse working environments. Following implications for organizations of job gender biases in postings, these may also require conditions that disproportionately affect one gender, such as long hours, that most impact women due to traditional caregiving roles [37]. Motherhood can lead to biased assessments of job applicants. Organizations need to recognize these biases and develop policies supporting working parents to create a diverse and inclusive work environment [56].

The role of cultural matching in hiring decisions is crucial to avoid gender biases and overcome limitations to diversity. By fostering a culture of inclusion and reducing the emphasis on cultural fit, organizations can promote diversity and ensure fairer hiring processes [57]. Exploring how gender stereotypes influence the hiring process in law firms, Gorman highlights the importance of organizational culture in overcoming gender biases and promoting diversity, counteract stereotypes, and creating an inclusive environment for all employees [58]. Kalev analyses the effectiveness of corporate affirmative action and diversity policies in promoting gender and racial diversity [59]. They suggest that organizations focus on structural changes, such as revising job postings and implementing diversity training, to create a more inclusive workplace culture.

The first hypothesis that using NLP techniques to identify and remove gendered language from job postings could result in a more diverse pool of candidates and ultimately lead to more women being hired for high decision-making positions, is, therefore, positive.

Though research manifests the need to continue overseeing potential gender biases in algorithms and NLP models, NLP techniques have been proven to recognize human biases in data [8]. Bohnet studied, for example, how behavioural design could be used to reduce gender biases in job postings [12]. She suggested that AI and NLP could replace gendered wording, improving more neutral job posting advertisements. Several authors believe that discussions between computer scientists, sociologists, and researchers in related fields may be the solution for improving latent gender bias found in machine learning data sets and model predictions. Multiple studies highlight the importance of interdisciplinary collaboration to develop methods to avoid biases in machine learning systems [59,60,61].

AI and NLP techniques offer significant potential to improve recruitment methods and reduce gender biases in high decision-making positions. By leveraging these techniques, recruiters can ensure that the best candidates, regardless of gender, are selected for senior positions, leading to more diverse and inclusive workplaces.

This research has showcased the potential of AI solutions to reduce discrimination within the selection process, regarding job postings. Research shows that “women are less likely than men to pass the initial screening phase” and that the management team, consciously or not, uses gender stereotypes when making recruitment decisions. Therefore, NLP techniques may be used, with the caution of a good definition of algorithms and the appropriate technique, to identify gender biases in big sets of information. They may also be used in past procedures to identify patterns, replying to the research question about the effectiveness of using NLP techniques to identify gender biases from job postings. It is, therefore, also useful to improve diversity in recruitment for high decision-making positions.

Since the distribution of both clusters shows a strong co-occurrence of terms associated with male roles, (agentic terms), another key finding of the research, is that the eventual heavier use of agentic terms or values is not valid to conclude that companies in more agentic sectors show a heavier use of agentic terms. The research has shown that when dealing with high decision-making positions, all sectors present a strong use of agentic terms, thus gendered language. The second hypothesis is, therefore, negative.

A robustness check in 100 FTSE companies can be done to test further hypothesis 2. Fig. 12 shows percentages of female presence on boards of the top 100 FTSE companies in 2022. We have sorted the results of the presence of women in high decision-making positions into the five sectors relevant to this study: IT/Tech, Transport, Financial, Banking, and Others. 27% of the top 100 companies of FTSE belong to these sectors.

The findings show that the most agentic sectors, “Banking”, “Finance”, and even “IT”, have more presence of women on boards and more women CEOs as the “Others sector”, and the “Transport” one, as regards women CEOs.

The better results, in terms of gender diversity in high decision-making positions of the traditional agentic sectors in the FTSE 100 companies of the UK, suggest that positive results of women's presence in high decision-making positions are linked to the culture in the corporations, as explained here above and other research, developing the used theories, has also evidenced.

Another limitation of the research may be the reduced number of sectors being considered. We have considered many documents, 2.198 grouped in just 5 sectors (IT/Tech; Transport; Financial S.; Banking; Others). New ways for future research could consider amplifying the number of sectors to validate the hypothesis.

Results suggest that real differences in gender equality in high decision-making positions lay in organizations that have undergone cultural transformations, not so much in the nature of the traditional agentic values of a sector. Bohnet emphasizes the importance of debiasing systems rather than individuals, which includes analysing job postings for gendered language, but also making necessary changes to create more inclusive and diverse workplaces [12].

By understanding and addressing the relationship between gender biases in job postings and the need for cultural transformations, organizations can create diverse and inclusive environments that benefit all employees and contribute to overall success.

On the other hand, non-binary genders [62] and other racial biases have scarcely been considered by NLP techniques applied to recruitment and should be considered in future work [63].

5. CONCLUSIONS

Is gendered language in job postings being used today? How may this gendered language prevent an equal representation of men and women in high decision-making positions? How are communal and agentic concepts and values present in job postings today? Our objective was to explore whether NLP techniques could be used to discover gender bias in job postings to explain possible reasons behind the underrepresentation of women in the corporate world.

Using an innovative experiment that extracts and analyses many job offers, this research has identified the most frequent terms and their rate of appearance in the advertisements, showing the gender bias they generated. The results show a strong co-occurrence of terms associated with male roles and the use of agentic-communal role differentiation. By acknowledging these roles and their influence on our perceptions, work can be done towards reducing gender biases.

With the contributions of different authors to the research of Social Theory, Think-Male-Think-Manager Theory, and Homosocial Theory, I have identified a table of terms used today relevant to analyse gender bias in job recruitment. These theories have been used to create a table for communal versus agentic language used in job postings for high decision-making positions today in the UK. Results showed that the use of NLP has been then helpful in identifying the selected terms in many jobs offers for discovering gender biases in recruitment for high decision-making positions**.**

Results have also shown that the most agentic sectors may not eventually present the strongest co-occurrence of agentic terms. They are certainly not showing the slowest percentage of gender diversity in decision-making positions.

Using gender neutral language in job postings emphasizes the diverse skills and qualities needed for the role, including those traditionally associated with women. Nevertheless, to change the underrepresentation of women in high decision-making positions, it is necessary to raise awareness of the critical importance of gender equality among the persons responsible for the selection processes. Moreover, developing policies that create a diverse and inclusive work environment is crucial. Improving recruitment methods is, therefore, a key strategy, which demands addressing cultural transformation in the organization in a transversal way, involving training, restructuration, changing cultural attitudes, creating diverse and inclusive environments that contribute to the overall success of the company, and the ability to attract female talent for high decision-making positions.

Ensuring fair hiring processes requires improving the culture of diversity and inclusion across the organization horizontally. This is a question of social welfare and a strategic and quality factor for corporations to create inner value.

COMPETING INTERESTS

Author has declared that no competing interests exist.

REFERENCES

1. Foley M, Williamson S. Does anonymising job applications reduce gender bias? Understanding managers' perspectives. Gender in Management. 2018;33(8):623-635.
2. Stephenson AL, Dzubinski LM, Diehl AB. A cross-industry comparison of how women leaders experience gender bias. Personnel Review; 2022.
3. Bhardwaj S. Decision-making in the recruitment of women on corporate boards: Does gender matter? Equality, Diversity and Inclusion. 2022;41(6):813-830. Available:https://doi.org/10.1108/EDI-08-2021-0188
4. Campion MC, Campion MA, Campion ED, Reider MH. Initial investigation into computer scoring of candidate essays for personnel selection. Journal of Applied Psychology. 2016;101(7):958-975. Available:https://doi.org/10.1037/apl0000108
5. Langer M, Landers RN. The future of artificial intelligence at work: A review on effects of decision automation and augmentation on workers targeted by algorithms and third-party observers. Computers in Human Behavior. 2021 ;123(106878). Available:https://doi.org/10.1016/j.chb.2021.106878
6. Gonzalez MF, Liu W, Shirase L, Tomczak DL, Lobbe CE, Justenhoven R, Martin NR. Allying with AI? Reactions toward human-based, AI/ML-based, and augmented hiring processes. Computers in Human Behavior. 2022;130(107179). Available:https://doi.org/10.1016/j.chb.2022.107179

7. Manning CD, Schütze H. Lexical acquisition. Foundations of statistical natural language processing. 1999; 999:296-305.
8. Raghavan M, Barocas S, Kleinberg J, Levy K. Mitigating bias in algorithmic hiring: Evaluating claims and practices. Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency. 2020;469-481.
9. Oswald FL, Behrend TS, Putka DJ, Sinar E. Big data in industrial-organizational psychology and human resource management: Forward progress for organizational research and practice. Annual Review of Organizational Psychology and Organizational Behavior. 2020;7(1):505–533. Available:https://doi.org/10.1146/annurev-orgpsych-032117-104553
10. Nissenbaum H, Freidman B. Bias in computer systems. ACM Trans. Inf. Syst. 1996;14(3):330–347. Available:https://doi.org/10.1145/230538.230561
11. Campolo A, Sanfilippo M, Whittaker M, Crawford K. AI Now 2017 Report. AI Now Institute; 2017.
12. Bohnet I. What works: Gender equality by design. Cambridge, MA: Belknap Press; 2016.
13. Bolukbasi T, Chang KW, Zou JY, Saligrama V, Kalai AT. Man is to computer programmer as woman is to homemaker? Debiasing Word Embeddings. Advances in neural information processing systems. 2016;29.
14. Zemel R, Wu Y, Swersky K, Pitassi T.amp; Dwork C. Learning fair representations. Proceedings of the 30th international conference on machine learning in proceedings of machine learning research. 2013;28(3):325-333. Available:https://proceedings.mlr.press/v28/zemel13.html
15. Hardt M, Price E, Srebro N. Equality of opportunity in supervised learning. Advances in neural information processing systems. 2016;29.
16. Garg N, Schiebinger L, Jurafsky D, Zou J. Word embeddings quantify 100 years of gender and ethnic stereotypes. Proceedings of the National Academy of Sciences. 2018;115(16):E3635-E3644.
17. Wang Y, Blei D, Cunninngham JP. Posterior collapse and latent variable non-identifiability. In M. Ranzato, A. Beygelzimer, Y. Dauphin, P.S. Liang & J. Wortman Vaughan, Advances in Neural Information Processing Systems 34. NeurIPS; 2021.
18. Carli LL. Gender, language, and influence. Journal of Personality and Social Psychology. 1990;59(5):941.
19. Lakoff R. Linguistic theory and the real world 1. Language Learning. 1975;25 (2):309-338.
20. Brownlow S, Rosamond JA, Parker JA. Gender-linked linguistic behavior in television interviews. Sex Roles: A Journal of Research. 2003;49(3-4):121–132. Available:https://doi.org/10.1023/A:1024404812972
21. Haas A. Male and female spoken language differences: Stereotypes and Evidence. Psychological Bulletin. 2007;86:616-626. Available:https://doi.org/10.1037/0033-2909.86.3.616
22. Leaper C, Ayres MM. A meta-analytic review of gender variations in adults' language use: Talkativeness, affiliative speech, and assertive speech. Personality and Social Psychology Review. 2007; 11(4):328–363. Available:https://doi.org/10.1177/1088868307302221
23. Newman ML, Groom CJ, Handelman LD, Pennebaker JW. Gender differences in language use: An analysis of 14,000 text samples. Discourse Processes. 2008;45(3):211–236. Available:https://doi.org/10.1080/01638530802073712
24. England P, Levine A, Mishel E. Progress toward gender equality in the United States has slowed or stalled. Proc. Natl. Acad. Sci. U.S.A. 2020;117:6990–6997. DOI: 10.1073/pnas.1918891117
25. Bertrand M. Gender in the twenty-first century. AEA Papers Proc. 2020;110:1–24. DOI: 10.1257/pandp.20201126
26. Ibarra H. Paving an alternative route: Gender differences in managerial networks. Social Psychology Quarterly. 1997;60(1):91-102.
27. Chattopadhyay P, Tluchowska M, George E. Identifying the ingroup: A closer look at the influence of demographic dissimilarity on employee social identity. Academy of Management Review. 2004;29(2):180-202.
28. Stout JG, Dasgupta N. When he doesn't mean you: Gender-exclusive language as ostracism. Personality and Social Psychology Bulletin. 2011;37(6):757-769.

29. Hu S, Al-Ani Jabir A, Hughes K, Denier N, Konnikov A, Ding L, Xie J, Hu Y, Tarafdar M, Jiang B, Kong L, Dai H. Balancing gender bias in job advertisements with text-level bias mitigation. Frontiers in Big Data; 2022.
Available:https://www.frontiersin.org/articles/10.3389/fdata.2022.805713
30. Born MP, Taris TW. The impact of the wording of employment advertisements on students' inclination to apply for a job. The Journal of Social Psychology. 2010;150 (5):485–502.
Available:https://doi.org/10.1080/00224540903365422
31. Eagly AH. Reporting sex differences. American psychologist. 1987;42(7):756-757.
Available:https://doi.org/10.1037/0003-066X.42.7.755
32. Cejka MA, Eagly AH. Gender-stereotypic images of occupations correspond to the sex segregation of employment. Personality and Social Psychology Bulletin. 1999;25(4):413-423.
33. Bem SL. Gender schema theory: A cognitive account of sex typing. Psychological Review. 1981;88(4):354-364.
34. Heilman ME. Description and prescription: How gender stereotypes prevent women's ascent up the organizational ladder. Journal of Social Issues. 2001;57(4):657-674.
35. Heilman ME, Okimoto TG. Why are women penalized for success at male tasks: The implied communality deficit. Journal of Applied Psychology. 2007;92 (1):81-92.
36. Rudman LA, Phelan JE. Backlash effects for disconfirming gender stereotypes in organizations. Research in Organizational Behavior. 2008;28:61-79.
37. Gaucher D, Friesen J, Kay AC. Evidence that gendered wording in job advertisements exists and sustains gender inequality. Journal of Personality and Social Psychology. 2011;101(1):109-128.
38. Fernandez-Mateo I, King Z. Anticipatory sorting and gender segregation in temporary employment. Management Science. 2011;57(6):989-1008.
39. Gupta D, Kumari S, Ekbal A, Bhattacharyya P. MMQA: A multi-domain multi-lingual question-answering framework for English and Hindi. In proceedings of the eleventh international conference on language resources and evaluation (LREC 2018), Miyazaki, Japan. European Language Resources Association (ELRA); 2018.
40. Kanter RM. Men and Women of the Corporation: New Edition. New York: Basic Books; 1993.
41. Schein VE. A global look at psychological barriers to women’s progress in management. Journal of Social Issues. 2001;57(4):675-688.
42. Ollilainen M, Calasanti TM. Metaphors at work: Maintaining the salience of gender in self-managing teams. Gender, Work & Organization. 2007;14(3):267-290.
43. Holgersson Ch. Recruiting managing directors: Doing homosociality. Gender, Work & Organization. 2013;20.
DOI: 10.1111/j.1468-0432.2012.00595.x
44. Schein VE. The relationship between sex role stereotypes and requisite management characteristics. Journal of Applied Psychology. 1973;57(2):95-100.
45. Gartzia L, Van Engen M. Are (male) leaders “feminine” enough? Gendered traits of identity as mediators of sex differences in leadership styles. Gender in Management: An International Journal; 2012.
46. Eagly Alice, Nater Christa, Miller David, Kaufmann, Michèle, Sczesny, Sabine. Gender stereotypes have changed: A cross-temporal meta-analysis of U.S. Public Opinion Polls From 1946 to 2018. American Psychologist. 2019;75.
DOI: 10.1037/amp0000494
47. NLTK. Documentation; 2023, January.
Available:https://www.nltk.org
48. Wu L, Hoi SC, Yu N. Semantics-preserving bag-of-words models and applications. IEEE Transactions on Image Processing. 2010;19(7):1908-1920.
49. Wordcloud 1.7.0.
Available:https://pypi.org/project/wordcloud/
50. Aizawa A. An information-theoretic perspective of tf–idf measures. Information Processing & Management. 2003;39(1):45-65.
51. Salton G, McGill M. Introduction to Modern Information Retrieval; 1983.
Available:https://towardsdatascience.com/text-summarization-using-tf-idf-e64a0644ace3
52. Ashnai B, Mani S, Kothandaraman P, Shekari S. Gender bias in the recruitment of entry-level B2B salespeople. Journal of

Business & Industrial Marketing. 2020;35(8):1335-1344. Available:https://doi.org/10.1108/JBIM-08-2019-0384
53. Quintero E. How are job applicants disadvantaged by gender based double standards in a natural setting. Cornell University; 2008.
54. Haslam SA, Ellemers N. Identity processes in organizations. Handbook of identity theory and research. 2011;715-744.
55. McDonald S. What's in the “old boys” network? Accessing social capital in gendered and racialized networks. Social Networks. 2011;33(4):317-330.
56. Correll SJ, Benard S, Paik I. Getting a job: Is there a motherhood penalty? American journal of sociology. 2007;112(5):1297-1338.
57. Rivera LA. Hiring as cultural matching: The case of elite professional service firms. American sociological review.32012;77(6):999-1022.
58. Gorman EH. Gender stereotypes, same-gender preferences, and organizational variation in the hiring of women: Evidence from law firms. American Sociological Review. 2005;70(4):702-728.
59. Kalev A, Dobbin F, Kelly E. Best practices or best guesses? Assessing the efficacy of corporate affirmative action and diversity policies. American Sociological Review. 2006;71(4):589-617.
60. Caliskan A, Bryson J, Narayanan A. Semantics derived automatically from language corpora contain human-like biases. Science. 2017;356(6334): 183-186.
61. West SM, Whittaker M, Crawford K. Discriminating systems: Gender, race, and power in AI. AI Now Institute; 2019.
62. Richards C, Bouman WP, Seal L, Barker MJ, Nieder TO, T’Sjoen G. Non-binary or genderqueer genders. International Review of Psychiatry. 2016;28(1):95-102.
63. Tang F, Kawamoto Y, Kato N, Liu J. Future intelligent and secure vehicular network toward 6G: Machine-Learning Approaches. Proceedings of the IEEE. 2020;108(2):292-307. Available:https://ieeexplore.ieee.org/document/8926369

Peer-review history:
The peer review history for this paper can be accessed here:
https://prh.ikprress.org/review-history/12023