Böhm, S., Linnyk, O, Kohl, J. Weber, T, Teetz, I., Bandurka, K. & Kersting, M. (2020). Analysing Gender Bias in IT Job Postings: A Pre-Study Based on Samples from the German Job Market. Proceedings of the 2020 on Computers and People Research Conference, p. 72-80.
https://doi.org/10.1145/3378539.3393862 https://dl.acm.org/doi/abs/10.1145/3378539.3393862

Analysing Gender Bias in IT Job Postings: A Pre-Study Based on Samples from the German Job Market

Stephan Böhm
RheinMain University
of Applied Sciences
Wiesbaden, Germany
stephan.boehm@hs-rm.de

Olena Linnyk*
Jens Kohl
Tim Weber
Ingolf Teetz
milch & zucker AG
Gießen, Germany
olena.linnyk@milchundzucker.de

Katarzyna Bandurka
Martin Kersting
Justus Liebig University of Gießen
Gießen, Germany
martin.kersting@psychol.uni-giessen.de

ABSTRACT

In Germany, as in many other industrial nations, there is currently a shortage of skilled workers in the IT sector, also known as the "war for talents". It is becoming increasingly difficult for companies to find suitable personnel using traditional recruiting instruments. Against this background, but also due to legal requirements, it is becoming more and more important that job postings are formulated in such a way that they have the greatest possible impact and no group of suitable applicants feels excluded. This study presents an approach that can be used to measure the gender bias in job postings in particular. A respective tool could provide recruiters with an instrument to identify and prevent unwanted gender bias. In our study, the prototype of such a tool will be developed and initially applied to analyse job postings in the IT sector of the German job market in comparison to samples from the automotive and health care sectors. We present some key statistics of this analysis and an outlook on future work.

CCS CONCEPTS

• **Social and professional topics → Employment issues**.

KEYWORDS

Gender bias; job postings; text analysis

ACM Reference Format:

Stephan Böhm, Olena Linnyk, Jens Kohl, Tim Weber, Ingolf Teetz, Katarzyna Bandurka, and Martin Kersting. 2020. Analysing Gender Bias in IT Job Postings: A Pre-Study Based on Samples from the German Job Market. In *Proceedings of the 2020 Computers and People Research Conference (SIGMIS-CPR '20), June 19–21, 2020, Nuremberg, Germany*. ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3378539.3393862

*Also with Frankfurt Institute for Advanced Studies, Frankfurt am Main, Germany.

SIGMIS-CPR '20, June 19–21, 2020, Nuremberg, Germany

ACM ISBN 978-1-4503-7130-8/20/06…$15.00
https://doi.org/10.1145/3378539.3393862

1 INTRODUCTION

The "war for talents" refers to the current competition for skilled workers in many industrial nations. In this it is becoming increasingly difficult for companies to find suitable personnel. In Germany, the competition for skilled workers will even increase in the future due to a predicted decrease in labour force [46].

In order to survive the war for talents, organisations try to find ways to become attractive employers for highly skilled job seekers. Unsurprisingly, also research deals with topics like organisational attractiveness, person-organisation (P-O) fit, corporate image, employer image or employer branding, to name but a few [16]. According to several studies it is very important that applicants perceive a match or sense of similarity between their personal and the organisational characteristics and values [7–9, 31] as cited in [16].

Job postings are an important recruiting instrument attracting the attention of qualified potential employees. A job posting is defined here as the announcement and dissemination of an open position by a company or other employers. The term is used in the following regardless of whether the job posting is published unpaid by the company itself (e.g., on the company's career website) or paid by a third party (e.g., a job portal provider) and thus also includes job advertisements. Job postings are one of the first ways of getting to know an organisation and allow job seekers to check if there is a match between themselves and the job or organisation. Finally, they can encourage potential employees to apply for the job [1]. Job postings were originally placed in print media but are now increasingly distributed online via the Internet. Consequently, job postings nowadays consist not only of text, but many other components such as logos, images, links to career websites, social media or even videos. Nevertheless, text remains the most important source for applicants to find out the job description, expected qualifications, requirements, tasks and benefits.

Against this background but also due to legal requirements concerning the equal treatment of all applicants, it is becoming more and more important that job postings are formulated in such a way that they have the greatest possible impact and no group of suitable applicants feels excluded. Nevertheless, recent research has shown that some job postings do not achieve this goal due to a hidden gender bias [20, 40]. In contrast to the explicit bias based on gender, the hidden gender bias reflects itself in a significant presence of gender stereotyped words in job postings. Gender stereotypes are generalizations and characterizations about the attributes and the behaviour of men and women [22]. They contain beliefs about how women and men typically are like (descriptive stereotypes) and beliefs about how men and women should be like (prescriptive stereotypes) [15, 22]. Research suggests that especially the presence of typically male-stereotyped words is problematic and hinders women from applying for a job [2, 20]. This not only leads to gender inequality, but also to a loss of potential talent through unintentional bias.

In this paper, we will examine which keywords contribute to a gender bias in job postings and how this gender bias can be measured. The present study is one of the first to focus on German-language texts, taking into account the specific characteristics of the German grammar (articles like *"der", "die", "das"*, in English *"the"*; gendered Endings of Nouns *"-er", "-in"*, etc.). A large corpus of German job postings is being used for the first time in order to do the statistical analysis and training of the machine learning models necessary to our investigations. In addition, a tool will be developed that automatically highlights words in job postings that might positively or negatively influence whether women feel addressed. The tool also calculates a gender bias score, which is suggested as a measure of the gender-neutrality of the job posting text and offers possible re-wordings to reduce a gender bias. Finally, we apply this tool to a random sample of job postings from the IT sector and other industries in order to check to what extent the terms identified in the keyword repository are relevant to the real job postings.

The paper is structured as follows: after the introduction, Section 2 describes the research background to gender bias in job postings. Section 3 presents related research and the research objectives of our study. Subsequently, Section 4 discusses how a keyword repository has been set up based on a literature survey of previous research on gender bias in job descriptions. That section also contains a proposal for the keyword-based calculation of a gender bias score in texts of job postings and elaborates on the development of a corresponding tool for the analysis of job postings. The approach and the findings of a brief study using the tool prototype for analysing a sample set of job postings from the German job portal jobstairs.de is contained in Section 5. Section 6 presents the discussion and conclusions followed by a section with an outlook on future research.

2 BACKGROUND ON GENDER BIAS

Gender bias and gender inequality are important actual research issues and closely related. Research on gender inequality in workplaces indicates that gender inequality can lead to not only moral but also economic consequences like financial losses and decreased innovation [10, 14] as cited in [43]. A meta-analysis [43] shows that perceived gender discrimination at work is negatively related to different personal outcomes (e.g., physical health outcomes and behaviors, psychological health outcomes) and to work-related outcomes (e.g., job attitudes, job-based work outcomes and relationship-based work outcomes).

Other studies examined gender bias in online contexts and systems. They found, e.g., that the user's gender information in Google Ads settings affects the advertisements that are shown to the user. In fact, Google showed less job postings for high-paying jobs to women than to men [11]. Further studies concentrated more on language style analyses, exploring, e.g., differences in words that are frequently used when writing about women or men. Results show an evidence for gender bias in Wikipedia where family- or gender-related words (e.g., *"husband", "women", "marriage", "child"*) are used more frequently in biographies of women than in biographies of men [44, 45].

Of growing importance is the occurrence of gender bias in word embeddings trained on free available online data (e.g., news articles, books, job postings websites). It has been found that word embeddings reinforce biases already present in the data, leading to analogies such as *"man is to computer programmer as woman is to homemaker"* [5]. Nevertheless, methodological options were also provided to reduce gender bias in word embeddings [5].

3 RELATED RESEARCH AND RESEARCH OBJECTIVE

Research has found subtle but systematic wording differences within job postings from traditionally male-dominated and female-dominated occupations [20, 40]. [20] analysed a total of 4,133 online job postings. The researchers created lists of stereotypical masculine and stereotypical feminine words and measured the total percentage of this keywords in the job postings. The results show that job postings within male-dominated areas contained greater masculine wording than job postings from female-dominated areas. There was no difference in feminine wording across the different areas.

[40] analysed about 17 million job postings collected over 10 years from the online platform LinkedIn. The researchers developed two scalable algorithms, comparable to metrics of the online services Textio and Unitive, that measure gender bias in job postings. They have found a gender bias towards masculine wording in the job market and identified the Technology sector as one of the most male-biased sectors [40].

Concerning the Information Technology (IT) sector, analyses showed an association between the IT profession and masculine gender stereotypes. 4,046 university students enrolled in IT courses in the U.S. were asked about their perceptions of the masculinity and femininity of IT skills. Results show that masculine gender stereotypes are being applied to the skills and knowledge that are required for the IT profession [42].

However, it is not only of interest whether a bias exists, but also what implications and consequences it has. Findings from several studies suggest that a gender bias in job postings can actively or passively discourage potential applicants from applying [2, 6, 20, 40]. The wording of job postings influences the accessibility of mental representations of women and men [36], the perceived gender diversity within the occupation and the job appeal [20]. Research indicates that male-biased wording in job postings *"perpetuates the image that these positions necessitate stereotypical male behaviour exclusively"* [23, p.732], and hence, reduces women's interest in these jobs and discourages them from applying [20]. The results indicate that women seem to be more influenced by gendered wording in job postings than men. Stereotypical male wording within job postings reduces the job appeal and the anticipated belongingness for women. Since women think that they will not belong, they may not apply for the job [20]. So stereotypical male-wording seems to *"push"* potential applicants away. Unbiased wording in job postings may therefore increase the number of applications. Furthermore, stereotypical female-wording might even *"pull"* potential applicants. These findings could not be replicated in another study [40], but it is important to mention that the methodology used was different (participants in this study evaluated mostly job postings from unfamiliar job areas).

Further results suggest that the type of linguistic description of, e.g., job requirements and skills, can also create a gender bias. Results indicate that women's willingness to apply is greater if job requirements are worded in a behavior-like way (verbs) than if they are worded in a trait-like way (nouns) [6, 47].

The main limitation of the literature on the effects of gender bias in job postings as far is that the studies are laboratory experiments, usually conducted on students. What is still missing is practice-oriented research on real job seekers. Another limitation of the mentioned literature is that the research was mainly conducted in English. To our knowledge, research on gender bias in German job postings concentrates mainly on job titles, comparing generic masculine forms vs. gender-fair forms [26, 27, 36]. These investigations indicate that the use of gender-fair forms, i.e., feminine-masculine word pairs (e.g., German *"Arbeitnehmerin und Arbeitnehmer"*, in Eng. *"employee, fem. and employee, masc"*.) or gender-unmarked forms (e.g., German *"Arbeitskraft"*, in Eng. *"employee"*) prevents a male bias, leads to a higher inclusion of women in mental representation [27], and might thus increase the probability of female applications. Further investigations and statistical analyses on German job posting texts are still missing.

The motivation of this paper is to fill this gap and to conduct practice-oriented research on German job postings. In contrast to English, the German language is a *grammatical gender language*. All nouns carry a grammatical gender (masculine, feminine or neutral) and other sentence parts related to a noun (such as articles or adjectives) change their form accordingly. Our research objective is to develop an instrument that identifies and prevents unwanted gender bias in German job postings, taking into account the specific characteristics of the German language and grammar. The tool should mark keywords relevant for gender bias and suggest improvements that reduce gender bias. The tool should also analyse the form of job titles and suggest gender-fair forms where necessary. Following previous research [20, 40], we will also analyse whether the identified keywords can be found in real job postings, and if so, to what extent. We have selected the IT sector as a sample for this analysis for two reasons: First, because it is particularly affected by the *"war for talents"*. And second, because previous English-language research has found it to be particularly affected by gender bias [40, 42].

The paper also intends to contribute to the further development of theoretical gender bias research. As mentioned above, gender bias has so far generally been investigated in relation to negative aspects that discourage women from applying for jobs. In our research, we will, therefore, also put those parts of job postings into focus that might foster women applying for jobs. For example, technical competencies and skills are often found in male-dominated occupational fields but are also necessary in job postings in order to be able to fill positions appropriately. If task-related words with push effect cannot be avoided, job postings may need to be enriched by other pull words that encourage women to apply. The distinction between such push and pull words can provide important impulses for further research if the corresponding word usage can be proven in practice.

4 APPROACH ON ANALYSING GENDER BIAS IN JOB POSTINGS

4.1 Defining the Keyword Repository

We identified lists of gender-stereotypical German words, frequently used in job postings. Using a corpus of approximately 500,000 job postings as collected on the *jobstairs.de* portal during the last ten years, we have built a gender detection model using deep learning. After training a custom vectorization algorithm (embedding), we followed the approach of [5] and defined gender bias or gender-stereotyping of a word as its projection onto the gender subspace of the word embedding, identifying the gender direction by gender pairs (e.g., she-he, her-his, woman-man, etc.). However, as noted also by [21, 39], this analysis alone failed to capture the full picture of gender bias. The authors of [21] suggest to use the notion of cluster bias, following the observation that small individual projections on to the bias axes can lead to a considerable bias in combination. For the identification of these clusters, it was, in turn, necessary to first identify male or female stereotypical words through human annotation.

In order to create a repository of the male or female stereotypical keywords, we reviewed existing English speaking literature on gender stereotypes generally [3, 12, 17, 32] and on the gender stereotypes in the wording of job postings particularly [6, 13, 20, 24, 41, 47]. Furthermore, we reviewed existing German speaking literature on gender stereotypes at the workspace [4, 29, 34, 37, 38] as well as on specific recommendations for German job postings [18, 25, 28, 35]. Cross-cultural studies suggest that there are more similarities between Western countries in gender stereotypes than differences [48]. Therefore, a part of our keyword repository is also based upon translations of the words from the English speaking research. Next, we have calculated the average correlation of these words with the gender axis of the embedding and found indeed that the male-stereotypical words have, on average, 56 percent larger correlation with the male axis than with the female in our embedding. Further more, the historical average of the job postings over the last 20 years shows higher prevalence of the male words in the postings for the male-stereotypical professions and vice versa (the statistics was limited by the small number of female-dominated professions). The language usage changes dynamically, but the historical data provide a good starting basis for our present study.

Our compiled keyword repository consists of three different parts. Firstly, it consists of a list of gender biased words that might discourage female potential applicants. This list contains stereotypical masculine adjectives and nouns or noun endings (e.g., *stärke, *gabe)[1] that are often used to describe job requirements (*"push"* words). Secondly, it includes another list of words that might encourage female potential applicants, which contains stereotypical feminine words, verbs or verb phrases that are often used to describe job requirements (*"pull"* words). Finally, we created a list of

[1] These German words are comparable to "strengths" and "abilities". However, in the German language, terms are often formed by stringing together several words. In the following, English terms will be used as examples. These words are only for understanding, but do not represent the German word groups used in our approach.

gender-unmarked words and phrases that can be used to replace gender biased words without changing the meaning should one aim to reformulate a given text.

4.2 Scoring and Measurement

We propose to quantify the gender bias in job postings by using a single number score that carries the information about the amount of gender biased words of both kinds ("push" and "pull"). The score S that we propose is calculated by the formula:

$$S = \frac{1}{1 + \exp(-x)}\ , \tag{1}$$

where

$$x = \frac{1}{N}\left(\sum_{\text{pull words}} w_i - \sum_{\text{push words}} w_i\right), \tag{2}$$

$$N = \left(\sum_{\text{pull words}} w_i + \sum_{\text{push words}} w_i\right), \tag{3}$$

and w is the weight, which is equal to 1 or 2. We take $w = 2$ for the words, which are established by more than one source (publication) as being gender-stereotypical and $w = 1$ if only one source identifies this word or the opinion is controversial in the reviewed literature. By construction, the score lies in the interval $(0, 1)$. For the neutral texts with zero gender stereotypical words the score is equal to 0.5. The higher the score, the more pull words are in the text. For the score below 0.5 there are more push than pull words. We use the logistic function in order to reflect the expectation that the effect of push words of the text sentiment perception by the applicant saturates at some level, in loose analogy with the reaction of neurological systems to stimuli generally [30].

4.3 BetterAds – Model Implementation for Analyzing Job Postings

We have developed an automated tool ("BetterAds") to analyse the text of job postings, which not only finds the gender-stereotypical words and calculates the score, but also provides suggestions for reformulating parts of the text in a more gender-neutral way.

The functionality consists of the following parts:

- The raw page html code is pre-processed to obtain the title and the main text of the job posting.
- The title of a job posting is analysed in order to extract job titles (e.g., *"Softwareentwickler"*, in Eng. *"software developer, masc."* or *"Putzfrau"*, in Eng. *"cleaning woman"*). The corresponding job title is highlighted in blue.
- A gender-neutral or gender-opposite term is proposed for the job title in a pop-in bubble (e.g., *"Softwareentwicklerin"*, in Eng. *"software developer, fem."* or *"Reinigungskraft"*, in Eng. *"cleaning person"*)
- The main text and the title are analysed to find words and expressions from the "push" and "pull" lists. These words and phrases are highlighted in color-coded manner.
- The BetterAd-score is calculated according to the formula (1) to indicate the gender-neutrality ($S \approx 0.5$) or the degree of bias ($S < 0.5$ for male-dominant, $S > 0.5$ for female-dominant) in the text.
- For each expression involving a push word, a suggestion is provided of how to reformulate it gender-neutrally or using pull words.

The tool has so far only been implemented as a prototype in order to be evaluated in a first test phase. An evaluation can focus on two important aspects in particular: Firstly, the question arises as to whether the identified keywords can be found at all in real job postings, and if so, to what extent. The second question is whether the word changes suggested by the tool can improve the way female applicants are addressed.

For the purpose of the second question, an ex-post analysis of job postings would be possible, but with restrictions. First of all, not only the postings would have to be analysed for the keywords, but also information on the gender of the applicants who have applied for the respective job postings has to be available. Since the performance of different job postings is compared, it would still remain unclear to what extent performance differences depend on the keywords or other factors such as the attractiveness of the jobs or employers. A better alternative is split testing, i.e., A/B testing, comparing the performance of the unchanged job postings with the optimised ones. However, as this article is work in progress, it will deal with the first question, i.e. the findability and frequency of the identified pull and push keywords in real job postings. Further research can be conducted in the future to answer the second question.

5 STUDY AND FINDINGS

Depending on the profession, the proportion of women and men can vary greatly. In the field of obstetrics and maternity care, for example, the proportion of women in Germany is 99.8 percent. At the other end of the spectrum, i.e. in jobs with a very low proportion of women, there are, for example, jobs in the construction industry, especially in facade construction. Here, the proportion of women is only 0.2 percent [19]. The reasons for these very different percentages of women are very diverse and can be due to social norms, salaries, physical requirements, existing employment structures or working environments. Even within one profession, the proportion of women can vary, for example, depending on whether they have regular employment contracts or management positions. In the short term, the existing distribution of men and women in an occupation is also likely to have a significant impact on the expected gender distribution of potential candidates as well as applicants, even though the distribution may change over time, for example because more men or women choose to study or train for a particular occupation.

Against this background, it is to be expected that gender bias in job postings also has a different significance depending on the occupational group. It is therefore plausible to assume that occupational groups in which only a few female applicants can be reached also formulate their job postings in a less female-oriented manner than is the case for applicant groups with a high proportion of women. In principle, however, the differences should not be too great, since employers are already required by law to refrain from gender-specific discrimination in job postings. However, on the other hand, if there is a chance of generating female applicants in a previously male-dominated profession, it can be very attractive for employers to intensify efforts to address women appropriately.

IT specialists are a male-dominated occupational group with a shortage of skilled workers and a high potential of qualified women [33]. In Germany, the proportion of women in IT and other ICT occupations is currently 16.5 percent [19]. The following study will therefore examine a sample set of IT job postings according to gender bias keywords. For comparison purposes, sample sets in occupational groups with high and low proportions of women will also be used. The procedure is described in the following section.

5.1 Scope of Research and Sample Selection

Within the framework of the study, job postings on the German job portal *jobstairs.de* were examined, with which the authors are conducting a joint research project. This portal is characterised by a special business model: companies do not place individual job ads, but agree on contract packages that allow for the posting of larger batches of postings. Therefore, the individual job categories are partly dominated by a few larger German companies that disseminate their job offers on this portal.

The selection of job postings was made at the end of January 2020. At that time there were over 32,000 active job postings in the portal. To select the comparison sets for the IT sector, the three job categories from the portal with the highest and lowest proportions of women employed in the industry were examined more closely. It should be noted that the occupational groups from the official job statistics could not directly be assigned to the portal's job categories. Thus, Table 1 shows the number of active job postings in the different categories as well as the the proportion of women employed according to the official job statistics [19] in the closest equivalent professional group. Based on these considerations, the portal's job categories Automotive and Automotive Suppliers and Health Care and Medical were selected as comparison groups as there were sufficient job postings from various companies in the respective categories.

Table 1: Selection of Job Categories for Comparison Sets

Portal's Job Category	Number of Jobs	Prop. of Women
Construction Industry	271	1.5%
Energy / Utility	18	3.9%
Automotive / Automotive Suppliers	6,612	7.7%
IT / IT Services	190	16.5%
Media	17	61.5%
Auditing / Tax Consultancy	1,320	74.9%
Health Care / Medical	493	82.2%

Source: The proportion of women employed in the industry was approximated by corresponding professional groups from [19].

Within each of the three groups IT/IT Services [ITS], Automotive/Automotive Suppliers [AAS] and Health Care/Medical [HCM], a random sample of 100 job postings was generated. These sets were refined by excluding non-German language ads and internship positions. This was required as the tool was realized based on German keywords only and internship postings are characterized by a different text structure. In a further step, the correct classification of the job categories was checked and job postings with incorrect classifications were eliminated manually.

5.2 Preliminary Findings in the Sample

The final sizes of the sample sets and some key statistics on the BetterAd scores are shown in Table 2. As mentioned before, a score of 0.5 represents a gender-neutral text of the job postings. Below 0.5, women tend to be discouraged from applying to a job by the dominance of push words. At scores above 0.5, the effects of the pull words that encourage women predominate. Looking at the statistical data, it can be seen that the mean BetterAd score was highest in job postings in the HCM group (M=0.4150, SD=0.1247), which has the highest proportion of women among employees. In job postings of the strongly male-dominated group AAS, the mean value of the BetterAd score was lowest (M=0.3241, SD=0.1197). In the area of job postings from the IT sector (ITS), the mean value (M=0.3948, SD=0.1011) of the score was between the aforementioned values, but also surprisingly high. A t-test for independent samples was performed between the groups of ITS and AAS as well as ITS and HCM to test the statistical significance of the observed deviations of the mean scores. The tests show: the differences of the score's mean value between ITS and AAS differ significantly ($t(126)=-3.509$, $p=0.001$), while there are no significant differences for IT and HCM ($t(143)=-1{,}073$, $p=0{,}285$).

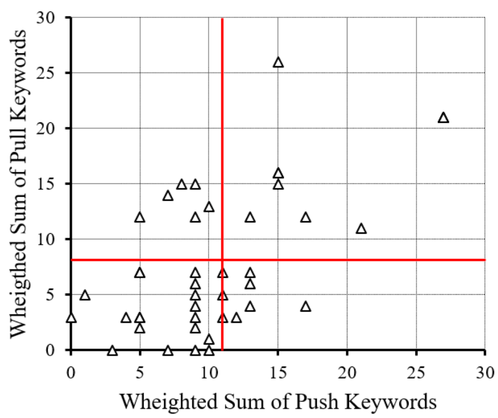
Table 2: Selection of Job Categories for Comparison Sets

Category	Param.	N	Min	Max	Mean	Std
AAS	Score	43	0.1192	0.7311	0.3241	0.1197
	Push2	43	0	13	4.77	3.146
	Push1	43	0	6	1.26	1.026
	Pull2	43	0	5	1.44	1.652
	Pull1	43	0	16	5.30	4.229
ITS	Score	85	0.1945	0.7711	0.3948	0.1011
	Push2	85	0	10	5.21	2.596
	Push1	85	0	9	2.56	2.107
	Pull2	85	0	6	2.01	1.842
	Pull1	85	2	25	9.22	4.199
HCM	Score	60	0.1720	0.7311	0.4150	0.1247
	Push2	60	0	7	1.43	1.609
	Push1	60	0	5	1.13	0.791
	Pull2	60	0	6	1.02	1.127
	Pull1	60	0	16	3.53	2.613

Interesting observations are also made on the frequencies of push and pull words. Table 2 shows separate word count statistics for keywords with single (Pull1, Push1) or double (Pull2, Push2) weighting in the score formula as described in Section 4.2. The mean value of Pull2 keywords encouraging applications from women is highest in the group of job postings ITS (M=2.01, SD=1.842) and lowest in HCM (M=1.02, SD=1.127). This could confirm the assumption that such words are used especially in those groups that are male-dominated but suffer from a lack of skilled workers and therefore try to recruit more women.

The differences in the use of push and pull keywords in the job posting categories can be illustrated using scatter plots. The following diagrams show the sums of the weighted number of push and pull keywords. Figure 1 shows this representation for the job postings from the AAS group. Each of this sums include, for example, the number of Push2 keywords multiplied by two as well as the value of the number of Push1 keywords for a job posting. In addition, the mean values of these sums are shown as red lines in the diagrams. Since AAS is a category with male-dominated professions, it is plausible that there are relatively many push keywords compared to pull keywords, which means a stronger gender bias exists that can discourage women from applying. It should be kept in mind that push keywords can be originated in the description of the job itself (e.g. *"analyst"*).

Figure 1: Keyword Plot for AAS



A different picture can be seen in Figure 2, which shows the values for the category HCM. Postings in this category are characterized by almost equally few pull and push keywords. This could be interpreted in a way that the job postings contain few keywords that are dissuasive or discouraging to women, but also that no special effort is made to make the job postings particularly attractive to women.

The wide range of values in the job posting category ITS in Figure 3 is interesting. In comparison to the other two categories, the highest average values are recorded for the totals of the push and pull keywords. In the case of the push keywords in this category of male-dominated professions, many are used to describe the specific requirements of the jobs, which are very technically and analytically oriented. Traditionally, recruiters describe these skills using the words associated with the male gender (e.g., "analyzing", "master"). However, the job postings in this category also show the highest mean values for the pull keywords. As suspected at the beginning, this could be an indication that employers are already making efforts to integrate more pull keywords in the intention to motivate women to apply.

Finally, we identify the most frequently observed Push2 and Pull2 keywords across our sample. The results are shown in Table

Figure 2: Keyword Plot for HCM

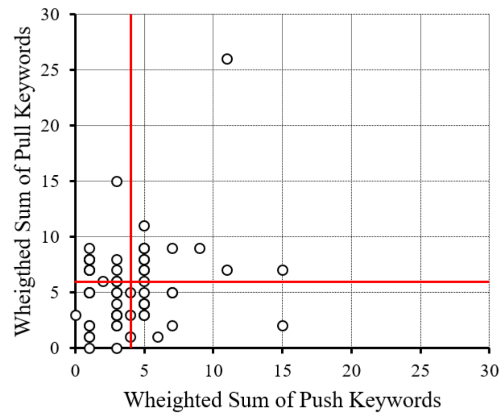
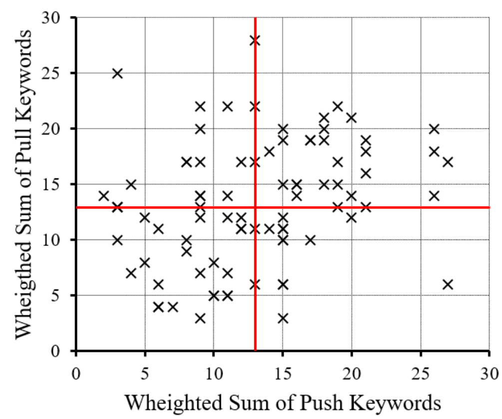


Figure 3: Keyword Plot for ITS



3. The listed terms represent the English translations of the word stems or word groups that occurred most frequently in the German-language analysis. The term *"ability"* thus stands for various terms that were explicitly demanded by the applicants in the job postings as qualifications (*"ability", "gift", "talent", "skill"*). Also shown are the proportions of corresponding keywords among the total number of keywords marked in the analysed job postings.

Table 3: Top 5 Word Stems for Push2 and Pull2 Keywords

Push2 Keyword	Prop.	Pull2 Keyword	Prop.
Ability	35%	Involvement	32%
Analysis	15%	Creativity	24%
Flexibility	9%	Communication	12%
Top/Best	4%	Motivation	11%
Autonomous	2%	Inspiration	9%

Based on the listed keyword stems, it is clear that the optimization of the corresponding job postings will offer some challenges.

For example, it is not enough to increase the number of keywords that encourage women to apply. It is also important to reduce masculine stereotypical keywords. As Table 3 shows, however, push keywords are often linked to the way in which knowledge and skills requirements are communicated, namely, in the static fashion: one either has *"analytical abilities", "communication talent"* and *"command of English language"* or not. Optimizing such job posting texts is a particular challenge, because it involves thoughtful reformulation into the skill requirements in the action form: *"you find solutions to difficult problems and readily understand complex structures", "you can communicate clearly"*. After all, a wording optimised job posting that appeals to women must nevertheless reflect the expected requirement profile as precisely as possible.

In the developed tool prototype, artificial intelligence methods were used to find suitable synonyms and suggested formulations that still adequately represent, for example, the knowledge requirements, but are better suited for female target groups. In further research work, it will be necessary to investigate not only to what extent job postings optimised in this way for the reduction of gender bias can indeed generate more female applicants, but if the text optimization does not discourage the male applicants, and also to what extent the applicants acquired due to the reformulation actually meet the job requirements.

6 DISCUSSION AND CONCLUSIONS

The objective of this research was to develop a prototypical tool to measure the gender bias of job postings. Furthermore, the tool should mark keywords relevant for gender bias and suggest changes to the wording, which improve gender-neutrality and possibly encourage a wider talent pool to apply. Based on the results of research on gender bias in the context of recruiting, a keyword repository of push and pull words was defined. Push words were considered to be those that in previous studies were proven to discourage women from applying, and pull words were those that tend to encourage women to apply. Moreover, a formula for the calculation of a gender bias (BetterAd) score was developed and implemented. Initial tests of the tool on three random sample sets of job postings from a German job portal showed that the tool delivers plausible results. Thus the proposed BetterAd score provides an orientation for identifying gender bias in job postings. However, the differences in the scores are small and not for all samples statistically significant, so that the validity of the model still needs to be further evaluated.

6.1 Practical Recommendations

An important finding of this study for practitioners can be found in the preparatory work for the development of the word repository. It does not appear to be advisable to rely solely on machine learning approaches for the identification of gender bias. Instead, expert-based approaches are preferable, which are also based on the current state of research. With the combination of expert knowledge and machine learning, however, approaches and tools can be developed which are applicable in practice.

The initial tests with our tool support the assumption that the number of push words in job postings of male-dominated occupational groups tends to be higher. The introduction of such tools can, therefore, sensitize the creators of job postings to avoid push words where possible and to replace them with gender-neutral formulations. The results of our study also indicate that pull words are used very differently across professional groups. Pull words are not necessarily being used to a greater extent where many women are applying anyway. Especially in the male-dominated IT sector, where there is a high shortage of skilled workers, pull words are already being used extensively. Companies in practice should, therefore, always keep these two dimensions of optimization in mind (substitution of push words and increased usage of pull words) when formulating job postings. For such an enrichment with pull words or a compensation of push words, tools like the proposed one can provide valuable assistance.

6.2 Theoretical Implications

This study is focused on applied research. Therefore, this paper was not intended to directly validate or extend theories in the field of gender bias research. However, the findings from the development and initial application of the tool may provide important impetus for gender bias research, particularly with regard to the proposed distinction between push and pull words. Previous research and theories have focused primarily on the negative impact of non-gender-neutral or masculine formulations in job postings. There is a lack of theoretical explanatory models of how beneficial pull words alone, but also in interrelation with push words, affect not only the perception of a job posting but also the intention to apply for jobs by women. The results of our initial tool deployment indicate that such compensation of push by pull words is already being attempted in practice in male-dominated occupational groups with a shortage of skilled workers and make the need for further research evident.

7 OUTLOOK ON FUTURE RESEARCH

This research is a work in progress. The approaches presented and the results of the study are therefore to be regarded as preliminary and are characterized by various limitations. The validation of the suggested score does not only require larger sample sizes, but also a more comprehensive set of job postings from a larger variety of employers. Future user tests with recruiters may not only investigate the validity of the BetterAd score but also the acceptance of wording improvements suggested by the tool. As already mentioned, split or A/B testing is necessary to find out whether the reduction of the gender bias can actually attract more female applicants. In addition, the fit of the applicants derived with the optimized job postings must be evaluated. Further limitations relate to the calculation of the score. In the existing formula, only the absolute word counts found in job postings have been considered. It should be examined to what extent the score provides robust results with regard to differing text lengths of the job postings. However, the tool has reached a stage of development that allows comprehensive testings within groups of selected lead users in practice. The findings of these tests will be reported in the near future.

8 ACKNOWLEDGMENTS

The study was carried out as part of the research project CATS (Chatbots in Applicant Tracking Systems) of the RheinMain University of Applied Sciences. This project (HA project no. 642/18-65) is funded in the framework of Hessen ModellProjekte, financed with funds of LOEWE – Landes-Offensive zur Entwicklung Wissenschaftlich-ökonomischer Exzellenz, Förderlinie 3: KMU-Verbundvorhaben (State Offensive for the Development of Scientific and Economic Excellence).

REFERENCES

[1] Kristin B Backhaus. 2004. An exploration of corporate recruitment descriptions on Monster.com. *The Journal of Business Communication (1973)* 41, 2 (2004), 115–136.

[2] Sandra L Bem and Daryl J Bem. 1973. Does Sex-biased Job Advertising "Aid and Abet" Sex Discrimination? 1. *Journal of Applied Social Psychology* 3, 1 (1973), 6–18.

[3] Iris Bohnet. 2016. In *Women Matter: Time to accelerate*. McKinsey. https://www.mckinsey.com/featured-insights/gender-equality/women-matter-ten-years-of-insights-on-gender-diversity/de-de

[4] Iris Bohnet. 2017. *What works: Wie Verhaltensdesign die Gleichstellung revolutionieren kann.* CH Beck.

[5] 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 *Neural Information Processing Systems (NIPS'16)*.

[6] Marise Ph Born and Toon W Taris. 2010. The impact of the wording of employment advertisements on students' inclination to apply for a job. *The Journal of social psychology* 150, 5 (2010), 485–502.

[7] Daniel M Cable and Jeffrey R Edwards. 2004. Complementary and supplementary fit: a theoretical and empirical integration. *Journal of applied psychology* 89, 5 (2004), 822–834.

[8] Daniel M Cable and Timothy A Judge. 1994. Pay preferences and job search decisions: A person-organization fit perspective. *Personnel psychology* 47, 2 (1994), 317–348.

[9] Daniel M Cable and Timothy A Judge. 1996. Person–organization fit, job choice decisions, and organizational entry. *Organizational behavior and human decision processes* 67, 3 (1996), 294–311.

[10] Jessica M Cornejo. 2007. An examination of the relationships among perceived gender discrimination, work motivation, and performance. *Electronic theses and dissertations. 3121* (2007).

[11] Amit Datta, Michael Carl Tschantz, and Anupam Datta. 2015. Automated experiments on ad privacy settings. *Proceedings on privacy enhancing technologies* 2015, 1 (2015), 92–112.

[12] Chris Dawson. 2020. What's in a Name: Professor of Practice - New faculty title draws industry experts to academia. Cornell University. https://www.engineering.cornell.edu/magazine/features/whats-name-professor-practice

[13] Dawn R Deeter-Schmelz, Andrea L Dixon, Robert C Erffmeyer, Kyoungmi Kim, Raj Agnihotri, Michael T Krush, and Ellen Bolman Pullins. 2018. Attracting Students to Sales Positions: The Case of Effective Salesperson Recruitment Ads. *Journal of Marketing Education* (2018). https://journals.sagepub.com/doi/10.1177/0273475318810335

[14] Robert L Dipboye and Adrienne Colella. 2005. The dilemmas of workplace discrimination. *Discrimination at work: The psychological and organizational bases* (2005), 425–462.

[15] Alice H Eagly and Steven J Karau. 2002. Role congruity theory of prejudice toward female leaders. *Psychological review* 109, 3 (2002), 573–598.

[16] Wim JL Elving, Jorinde JC Westhoff, Kelta Meeusen, and Jan-Willem Schoonderbeek. 2013. The war for talent? The relevance of employer branding in job advertisements for becoming an employer of choice. *Journal of Brand Management* 20, 5 (2013), 355–373.

[17] Katherine TU Emerson and Mary C Murphy. 2015. A company I can trust? Organizational lay theories moderate stereotype threat for women. *Personality and Social Psychology Bulletin* 41, 2 (2015), 295–307.

[18] Marianne Helfer Herrera Erazo. 2008. *Kompetente Bewerberinnen und Bewerber finden. Tipps und Tricks für die Gleichbehandlung von Frauen und Männern in Stelleninseraten.* Gleichstellungsbüro des Kantons Basel-Stadt (Hg.) - Basel. http://www.bs.ch/publikationen/gleichstellung/kompetente-bewerberinnen-und-bewerber-finden.html

[19] Bundesargentur für Arbeit. 2019. *Beschäftigte nach Berufen (Klassifikation der Berufe 2010) - Deutschland, West/Ost und Länder (Quartalszahlen) - Juni 2019.* https://statistik.arbeitsagentur.de/Statistikdaten/Detail/201906/iiia6/beschaeftigung-sozbe-bo-heft/bo-heft-d-0-201906-xlsx.xlsx

[20] 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), 109–128.

[21] 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. In *North American Chapter of the Association for Computational Linguistics (NAACL'19))*.

[22] Madeline E Heilman. 2012. Gender stereotypes and workplace bias. *Research in Organizational Behavior* 32 (2012), 113–135.

[23] Madeline E Heilman and Suzette Caleo. 2018. Combatting gender discrimination: A lack of fit framework. *Group Processes & Intergroup Relations* 21, 5 (2018), 725–744.

[24] Tanja Hentschel, Susanne Braun, Claudia Verena Peus, and Dieter Frey. 2014. Wording of advertisements influences women's intention to apply for career opportunities. In *Academy of Management Proceedings*, Vol. 2014. Academy of Management Briarcliff Manor, NY 10510, 15994.

[25] Tanja Hentschel and L. K. Horvath. 2015. Passende Talente ansprechen - Rekrutierung und Gestaltung von Stellenanzeigen. In *C. Peus, S. Braun, T. Hentschel & D. Frey (Hrg.), Personalauswahl in der Wissenschaft - Evidenzbasierte Methoden und Impulse für die Praxis*. Springer-Verlag, 65–82.

[26] Lea Hodel, Magdalena Formanowicz, Sabine Sczesny, Jana Valdrová, and Lisa von Stockhausen. 2017. Gender-fair language in job advertisements: A cross-linguistic and cross-cultural analysis. *Journal of Cross-Cultural Psychology* 48, 3 (2017), 384–401.

[27] Lisa K Horvath, Elisa F Merkel, Anne Maass, and Sabine Sczesny. 2016. Does gender-fair language pay off? The social perception of professions from a cross-linguistic perspective. *Frontiers in psychology* 6 (2016), 2018.

[28] H. W. Jablonski, T. Hentschel, and E. T. Neuhaus. 2015. Wieder nicht alle erreicht? Diversity Wording für eine effiziente Ansprache vielfältiger Zielgruppen. In *Diversity Konferenz*. Charta der Vielfalt e. V., Der Tagesspiegel.

[29] Uwe Peter Kanning. 2015. *Personalauswahl zwischen Anspruch und Wirklichkeit: eine wirtschaftspsychologische Analyse*. Springer-Verlag.

[30] Kenneth Knoblauch and Laurence T. Maloney. 2012. *The Psychometric Function: Introduction*. Springer New York, New York, NY, 107–139. https://doi.org/10.1007/978-1-4614-4475-6_4

[31] Amy L Kristof. 1996. Person-organization fit: An integrative review of its conceptualizations, measurement, and implications. *Personnel psychology* 49, 1 (1996), 1–49.

[32] Sarah-Jane Leslie, Andrei Cimpian, Meredith Meyer, and Edward Freeland. 2015. Expectations of brilliance underlie gender distributions across academic disciplines. *Science* 347, 6219 (2015), 262–265.

[33] Lydia Malin, Anika Jansen, Susanne Seyda, and Regina Flake. 2019. *Fachkräfteengpässe in Unternehmen Fachkräftesicherung in Deutschland – diese Potenziale gibt es noch*. Institut der deutschen Wirtschaft Köln e.V. https://www.kofa.de/service/publikationen/detailseite/news/kofa-studie-22019-fachkraefteengpaesse-in-unternehmen

[34] Claudia Peus, Susanne Braun, Tanja Hentschel, and Dieter Frey. 2015. *Personalauswahl in der Wissenschaft - Evidenzbasierte Methoden und Impulse für die Praxis*. Springer.

[35] Paula Risius, Zuzana Blazek, Anna Schopen, Sibylle Stippler, Sabine Brinkmann, Elena Reifenröther, Filiz Karsligil, Anja Seng, and Lana Kohnen. 2018. *Mit Stellenanzeigen gezielt weibliche Fachkräfte gewinnen*. Institut der deutschen Wirtschaft Köln e.V. https://www.kofa.de/mitarbeiter-finden-und-binden/mitarbeiter-finden/wen-rekrutieren/frauen

[36] Sabine Sczesny, Magda Formanowicz, and Franziska Moser. 2016. Can gender-fair language reduce gender stereotyping and discrimination? *Frontiers in Psychology* 7 (2016), 25.

[37] Rainer Spies. 2012. Hartnäckige Stereotype im Spiegel der Wissenschaft, Das Thema Frauen und Karriere auf dem Herbstworkshop der Kommission Personalwesen. In *Personalführung*, Vol. 11. 56–59.

[38] Melanie Steffens and Irena D Ebert. 2016. *Frauen–Männer–Karrieren: Eine sozialpsychologische Perspektive auf Frauen in männlich geprägten Arbeitskontexten*. Springer-Verlag.

[39] Tony Sun, Andrew Gaut, Shirlyn Tang, Yuxin Huang, Mai ElSherief, Jieyu Zhao, Diba Mirza, Elizabeth Belding, Kai-Wei Chang, and William Yang Wang. 2019. Mitigating Gender Bias in Natural Language Processing: Literature Review. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics, Florence, Italy, 1630–1640. https://doi.org/10.18653/v1/P19-1159

[40] Shiliang Tang, Xinyi Zhang, Jenna Cryan, Miriam J Metzger, Haitao Zheng, and Ben Y Zhao. 2017. Gender bias in the job market: A longitudinal analysis. *Proceedings of the ACM on Human-Computer Interaction* 1, CSCW (2017), 1–19.

[41] Toon W Taris and Inge A Bok. 1998. On gender specificity of person characteristics in personnel advertisements: A study among future applicants. *The Journal of psychology* 132, 6 (1998), 593–610.

[42] Eileen M Trauth, Curtis C Cain, Kshiti D Joshi, Lynette Kvasny, and Kayla M Booth. 2016. The influence of gender-ethnic intersectionality on gender stereotypes about IT skills and knowledge. *ACM SIGMIS Database: the DATABASE for Advances in Information Systems* 47, 3 (2016), 9–39.

[43] María del Carmen Triana, Mevan Jayasinghe, Jenna R Pieper, Dora María Delgado, and Mingxiang Li. 2019. Perceived workplace gender discrimination and employee consequences: a meta-analysis and complementary studies considering country context. *Journal of management* 45, 6 (2019), 2419–2447.

[44] Claudia Wagner, David Garcia, Mohsen Jadidi, and Markus Strohmaier. 2015. It's a man's Wikipedia? Assessing gender inequality in an online encyclopedia. In

Ninth international AAAI conference on web and social media.
[45] Claudia Wagner, Eduardo Graells-Garrido, David Garcia, and Filippo Menczer. 2016. Women through the glass ceiling: gender asymmetries in Wikipedia. *EPJ Data Science* 5, 1 (2016), 5.
[46] Martin Werding. 2019. Talente werden knapp: Perspektiven für den Arbeitsmarkt. In *War for Talents*. Springer, 3–17.
[47] Lien Wille and Eva Derous. 2018. When job ads turn you down: how requirements in job ads may stop instead of attract highly qualified women. *Sex Roles* 79, 7-8 (2018), 464–475.
[48] John E Williams and Deborah L Best. 1990. *Measuring sex stereotypes: A multination study, Rev.* Sage Publications, Inc.