

Gendered language and employment Web sites: How search algorithms can cause allocative harm

by Karin van Es, Daniel Everts, and Iris Muis

Abstract

Research on algorithms and artificial intelligence in the hiring process tends to focus on applicant screening and is often centered on the employer perspective. The role played by intermediaries, such as employment Web sites in the distribution of information about employment opportunities, tends to be overlooked. This paper examines the role of search algorithms on employment Web sites and their retrieval of employment opportunities for job seekers based on gendered search terms. Through a basic algorithm audit of the search engines used by three major employment Web sites active in the Dutch job market, we explore whether their search algorithms neutralize or reinforce gendered language, in case of the latter thereby naturalizing stigmas and pre-existing bias. According to our findings, employment Web sites can cause allocative harm if they repeatedly fail to show all opportunities relevant to job seekers.

Contents

[Introduction](#)

[Employment Web sites: Algorithmic bias and discrimination](#)

[Gendered language and job applications](#)

[Gendered language, search algorithms and the allocation of employment opportunities](#)

[A basic algorithm audit](#)

[The neutralization of gendered terms](#)

[A quantitative perspective](#)

[Conclusion](#)

Introduction

Biases, understood as inclinations in favor or against something or someone, permeate society. For a long time, algorithms have been seen as a remedy for human bias, under the assumption that computers simply compute what they have been programmed to compute [1]. However, algorithms are now more widely understood as not neutral (e.g., Crawford, 2017; Gillespie, 2014; Noble, 2018). While *discriminating between* factors is inherent to algorithmic systems and could establish particular kinds of bias, it may also lead to the systematic and unfair *discrimination against* certain groups or individuals (Wieringa, 2021), meaning these groups or individuals are then denied equal rights and opportunities. It is therefore important then that we strive to understand the operations and impact of algorithms (Gillespie, 2014).

Currently, within the job market, algorithms are widely deployed in support of the hiring process. Here we can distinguish between roughly four phases: sourcing, screening, interviewing and selection (Bogen and Rieke, 2018). During sourcing, a candidate pool is built, and this group of potential hires is subsequently screened for suitable interview candidates. At the end of the hiring process, one or more of the interviewees are actually offered a position (Chen, *et al.*, 2018).

Research on algorithms and artificial intelligence in the hiring process tends to focus on the recruiter's or employer's perspective (e.g., Ajunwa and Green, 2019; Borgen and Rieke, 2018; Chen, *et al.*, 2018). With regard to algorithmic bias, screening — how algorithms score and sort resumes to rank applicants — has attracted considerable attention (e.g., Raghavan, *et al.*, 2020).

Overlooked in research has been the role of tech intermediaries in the distribution of information about employment opportunities (Kim, 2020). To address this gap, in this paper we pay attention to a crucial process linked to sourcing: the role played by search algorithms on employment Web sites. Specifically, we are concerned with the process that *job seekers* go through when searching for job openings on employment Web sites.

Potential bias and discrimination in the online job market in the Netherlands (and other countries) are complicated by gender-inflected language. The Dutch language has feminine and masculine terms for many occupations and roles. In some cases, there is an umbrella term covering both genders (Pous, 2020). Many male-gendered occupational pronouns can be used to refer to women because, grammatically, they are generically masculine, but this does not work the other way around. The use of gendered search terms may affect the visibility of job openings and exclude job seekers from opportunities. In this paper, we focus on the exclusionary impact of search on employment Web sites in relation to gendered language. The following research question will be addressed: how do the search algorithms used on Dutch employment Web sites deal with gendered language?

The paper's first part explores algorithms and their relation to bias and discrimination within the context of job recruitment. We subsequently zoom in on the relationships between algorithms and gendered language. We then discuss the algorithm audit of three employment Web sites active on the Dutch job market and the results of this research. Here, it becomes apparent that the search engines of employment Web sites can deprive job seekers of employment opportunities by normalizing gendered job titles. As we will elaborate upon, in these instances, a representational harm is transformed into an allocative harm (Crawford, 2017), because certain employment opportunities have been withheld from the job seekers' view. In conclusion, we underscore the need to study the role of intermediaries in the distribution of information, and outline the responsibility employment Web sites should take on so as to counter the bias resulting from gendered language.

Employment Web sites: Algorithmic bias and discrimination

In the face of a long-held popular belief, algorithms are not neutral [2]. They can reproduce as well as enhance biases. In light of differences between computer science and the humanities in the use of certain terms, we find it important to reiterate that we use bias here to refer to “slant” or “inclination.” Moreover, we provide that algorithmic systems discriminate *between* factors and have the potential to systematically and unfairly discriminate against people.

According to Friedman and Nissenbaum (1996), bias can be a part of algorithmic systems in three ways. First, bias in algorithms can result from already present, societal bias rooted in institutions, practices and attitudes — here, we can speak of *pre-existing bias*. Second, bias can result from technical design constraints — *technical bias*. Third, *emergent bias* occurs when an algorithm is used in unexpected contexts or by audiences different from those intended by their creators; bias then literally emerges, as it were, in an algorithm's usage. In this research, we concentrate primarily on pre-existing bias and the way search algorithms deal with it, although we should note that the three forms of bias are closely related and can each be seen as extensions of the others; in reality, these bias types often blend into one another.

To understand how pre-existing bias bears on how algorithms function, it is important to consider that algorithms are, in essence, classifications of reality. Algorithms reduce real-world complexity into a set of abstract instructions on how to deal with input coming from a messier, more nuanced reality. Simplifying reality in this way involves selection, reduction and categorization [3]. Or, as Gitelman and Jackson [4] put it, “Data need to be imagined as data to exist and function as such, and the imagination of data entails an interpretive base.” Developers must choose which variables are relevant for the task at hand and which features of an algorithm should be optimized. These design decisions are informed by the worldviews of programmers or other actors involved in the development of algorithms and reflect the particular socio-economic and historic context from which they emerge.

Programmers might, for instance, consider the use of gendered language important in deciding which job vacancies to show based on specific search terms. They might find that the use of a female-gendered pronoun should lead only to female-gendered job vacancies. Alternatively, programmers who, for instance, have never experienced any form of sexism might not deem gendered language an important factor at all. These programmers might create a search algorithm

that, when faced with a gendered part of speech, only shows job vacancies that use that same part of speech. In both cases, the end result would be a gender-biased search algorithm. Thus software developers' assumptions about how reality works are often incorporated into algorithmic functions, which then turn out to be biased and might even be discriminatory once they are put into use [5].

Aside from bias through the design of algorithms, bias can also be introduced to algorithms through the datasets on the basis of which they are developed. Computer and data scientists discover patterns in relevant datasets and, based on these patterns, create a picture of reality translated into computer code [6]. Datasets, however, are likely to reflect the generally unequal social, historical and economic conditions across society. Thus, algorithmic models based on them run the risk of reiterating biases present in the data itself [7].

In addition to *reproducing* social inequality, algorithms can also strengthen or *reinforce* biases in society due to their scale or ubiquity (O'Neil, 2016; Noble, 2018). For example, an algorithm that consistently puts a certain group of people at a disadvantage, makes them less likely to qualify for a job. As a result, fewer individuals from that particular group wind up doing that particular job, and this outcome will then be reflected in a subsequent employment dataset. If that dataset is used to develop the next version of the algorithm, it will be even more likely to conclude that the particular disfavored group is unsuitable for employment [8].

A case in point is a past version of Amazon's company-developed algorithm for selecting job seekers' resumes, which had not been ranking resumes for developer jobs and technical positions in a gender-neutral fashion. The algorithm put the resumes of female applicants at the 'bottom of the pile', because it had taught itself — training on resumes submitted to the company over the course of 10 years — that male candidates were more desirable (Dastin, 2018).

Bias in algorithms can also be more directly caused by pre-existing bias through the usage of algorithmic systems. In the past, this tendency has been clearest in relation to self-learning algorithms, which can become increasingly biased over time. Research by Chen, *et al.* (2018) has shown that employers using internationally operating job Web sites Indeed, Monster, and CareerBuilder would more often click on the curricula vitae of men than those of women — presumably a manifestation of the prejudice that men are generally better suited for jobs than women. Eventually, it was found that this caused the search algorithms used on those Web sites to learn that men's curricula vitae were preferable to those of women. It established a vicious circle: men's curricula vitae, more often put forward, reinforced the pre-existing prejudice that men are more suitable for employment than women (Chen, *et al.*, 2018).

Similarly, and relevant to our research, algorithms have been known to contribute to the normalization of gender stereotypes in language use. For example, the Google translation engine has exhibited bias in its translation of the Turkish gender-neutral phrases 'he/she/it is an engineer' and 'he/she/it is a cook', which it translated to 'he is an engineer' and 'she is a cook' (Morse, 2017). In this way, algorithms help to naturalize and amplify traditional stereotypes, such as, in this case, that technical professions are best suited to men and women are better suited for work in the kitchen.

To be clear, it is not a matter of *if*, but a matter of *to what degree* and *with what effects* algorithms are biased. Our aim here is to consider how bias expresses itself through search engines on employment Web sites and with what consequences. This is an important task to take on, because instances of harm can occur.

Kate Crawford (2017), principal researcher at Microsoft Research, distinguishes two types of harm: representational harm and allocative harm. Representational harm takes place in instances when a system reinforces stereotypes, whereas allocative harm occurs when a system unfairly withholds resources and/or opportunities, or allocates them unfairly. Google's translation engine, for instance, causes representational harm when it reproduces gender stereotypes and perpetuates sexism. Allocative harm is inflicted when, for example, a bank's algorithmic system denies mortgages to people living in certain U.S. zip codes.

As mentioned, bias can lead to the systematic and *unjust discrimination* against individuals or groups. Discrimination can occur both directly and indirectly. Direct discrimination occurs when a person is treated differently on the basis of a legally protected characteristic such as gender, race, and/or age [9]. These characteristics are known as 'protected criteria', and groups defined on their basis are generally referred to as 'protected groups' or 'protected classes' [10]. Indirect discrimination occurs when the treatment of an individual is based on characteristics that, while not protected by law, correlate with protected criteria in such a way that individuals from a protected class are nonetheless treated in a significantly different way than others [11].

Recent research has revealed that several Dutch employment platforms enable employers to indicate discriminatory preferences in their job advertisements, even though these Web sites' terms of service explicitly prohibit discriminatory postings. The problem exists because these sites rarely check the content of newly posted vacancies before they are published. The same research also found that employers are able to filter the curricula vitae of job applicants on the basis

of protected criteria (Inspectie SZW, 2021). Eight employment sites active on the Dutch market were researched. Two of those provided employers with the ability to differentiate potential candidates on the basis of nationality and age. One of them also allowed employers to select candidates on the basis of gender. Selecting candidates on the basis of any of these characteristics is prohibited by Dutch law.

Gendered language and job applications

A particular consideration in search engines on job application Web sites regarding bias and harm is its dealings with gendered language. Unlike English, Dutch is a gender-inflected language: most nouns are masculine or feminine. In line with the central role that males, in particular white, heterosexual and, we should add, cisgender men, traditionally occupy in Western society [12], terms such as ‘doctor’ and ‘poet’ — which originated as generic masculine nouns — are now often considered the ‘norm’. That is, they are used in a ‘neutral’ fashion to refer to any individual with such a position. Yet nouns such as ‘doctress’ and ‘poetess’ terms that in English are now completely discarded — can only refer specifically to women, underscoring the male as norm [13].

The recent trend in Dutch newspapers has been to avoid female pronouns entirely in favor of gender-neutral language in a bid for ‘fairer’ linguistic usage. Critics remark, however, that this practice only further cements the male as the norm and renders the women in these positions invisible. Here, little is done to combat inequality and prejudice (Pous, 2020).

In practice, the normative position of men and of masculine nouns and pronouns exercises adverse effects on individuals who do not fit the traditional description of the (white, heterosexual, cisgender) man. For example, in relation to the labor market, strongly gendered language has been found to affect how potential future employers evaluate job applicants: women benefit from using masculine job titles and are in fact devalued when using feminine job titles (Formanowicz, *et al.*, 2013). Moreover, it has been found that women are less likely to apply for jobs that are advertised with generic masculine pronouns. Additionally, women do not perform as well in interviews for these jobs (Stout and Dasgupta, 2011).

Linguistic usage also affects how job seekers perceive job openings. For example, research has shown that women are relatively more likely to respond to job postings in which both a male and female-gendered occupational title is mentioned — *e.g.*, “chairman/chairwoman” — than to job postings in which only a male-gendered occupational title is used as a neutral term — *e.g.*, “chair (m/f)” (Pous, 2020). Furthermore, it has been demonstrated that men and women alike are more likely to feel that they can be considered for a job that mentions male and female-gendered occupational titles (Chatard, *et al.*, 2005; Vervecken, 2013). This, in part, can be explained by what is here a reinforcement, via the use of both male and female occupational titles, of the idea that women can be as successful in traditionally male occupational practices as their male counterparts (Stahlberg and Sczesny, 2001; Vervecken, 2013).

In short, it matters how job applicants choose to describe themselves, as well as how job openings are described. These choices, however, are not independent of outside influence: people’s actions are informed by the way people as well as more abstract issues are already spoken about, thought of and dealt with within a given culture [14]. To put it more concretely, if it is ‘normal’ or ‘natural’ to associate certain positions with women, then job openings for that position are also more likely to be described with the use of female-gendered occupational titles. At the same time, women would be more inclined to see themselves in those jobs, to look for those jobs and, when they apply for those jobs, to describe themselves as women — that is, through the use of female-gendered language.

This type of naturalization is generally understood as the result of a constant repetition of a particular brand of linguistic usage, without significant contradiction [15]. It is here — when search algorithms simply take and re-produce pre-existing biases without doing anything to contradict gendered language that privileges the male norm, or when they hide the female nouns and pronouns — that they can cause representational harm.

Gendered language, search algorithms and the allocation of employment opportunities

The harms of algorithms in relation to gender biases extend further than representational harm. Before curricula vitae can be processed by an algorithm, job seekers must first choose to apply for a job — a choice that essentially determines which people, out of the job-seeking population, can become part of a job’s applicant pool (Peng, *et al.*, 2019). After all, a person can apply for an advertised position only if they actually know it is available. Algorithms participate in the

distribution of information by showing ‘relevant’ search results in a specific order, making some opportunities more visible than others and potentially obscuring some completely (Kim, 2020). It is here that we identify a form of potential allocative harm.

Indeed, both in the past and in the present, the varying visibility of opportunities for different groups has been identified as an issue. For instance, Facebook’s targeted advertising functionalities, which enabled companies to exclude women and older workers from exposure to specific job advertisements and thus blocked their access to employment opportunities, has previously been criticized on the grounds that it violated anti-discrimination laws (Tobin, 2019). Similarly, Google was accused of discriminatory advertisement delivery practices when researchers found that highly paid jobs were advertised less frequently to women (Datta, *et al.*, 2015).

This problem persists even today: in 2021, Facebook was once again discovered to have exhibited gender bias in its delivery of job advertisements (Basileal, *et al.*, 2021). Tackling these types of algorithmic issues poses unique challenges, because people often do not know they have been deprived of opportunities, while anti-discrimination laws in the U.S. that could help combat such prejudiced algorithms are primarily complaint-driven (Wachter, 2020).

Allocative harm, as described above, does not occur spontaneously. Rather, we find that employment Web sites can *transform* representational harm into allocative harm, because their search algorithms allocate access to employment opportunities (Kim, 2020). For instance, when algorithms on employment Web sites reinforce gender stereotypes by only relating a particular gender to certain employment opportunities (representational harm), the result might be — depending on how the algorithms in question function — that those using only female-gendered pronouns will only receive search results specifically related to female-gendered job applications. In this case Web sites would actually be preventing specific job seekers from seeing particular employment opportunities simply because of their use of gendered language (allocative harm).

Ideally, search algorithms on employment Web sites counteract societal, pre-existing bias in gendered language by neutralizing it — that is, by presenting all users with the same search results, despite their use of either gendered or gender-neutral language. It helps to ensure equal opportunities and mitigates allocative harm. Whether a search algorithm counteracts or reinforces pre-existing bias depends on its technical operation.

From a preliminary standpoint, we foresee three possible ways the search algorithms examined in this paper might function: they might make use of a word-stem taxonomy, or semantic scoring of proximity, or one-to-one matching. Word-stem taxonomies work as follows: when a job seeker queries [cook] the search algorithm retrieves all job postings containing that term. A search algorithm could be based on word roots and word families. In that case the core of the algorithm is a taxonomy in which the roots of all possible keywords form the core. Input search terms are then related to the root of the queried term, which the algorithm extracts from its taxonomy. For example, when searching with [cook], [cooking] or [chef], these search terms are all related to the root word [cook]. The results shown will be the same for all terms that relate to the same root (Sánchez and Moreno, 2004).

The semantic scoring of proximity works through the use of word embeddings: a technique more sophisticated than those based in word stems and word families. Word embeddings give words/language a mathematical representation in hundreds or thousands of dimensions, enabling ‘real’ semantic similarity to be found. This process goes one step further than comparison with word roots (Sánchez and Moreno, 2004). To illustrate: the search queries [a colleague] and [someone you often sit with in the office] would receive semantically similar scores, whereas a system using a word-stem taxonomy would not consider these two search queries to be related.

Finally, a search algorithm may also search purely on the basis of the exact term directly entered in the search bar, in which case we can speak of one-to-one matching. Obviously, this is the least complex kind of search algorithm. As one can imagine, all these various possible workings of search algorithms affect how gender stereotypes are reflected on employment Web sites when one performs search queries.

Certain types of algorithms handle gendered language differently than others; a search algorithm based on word-stem taxonomies will probably be better equipped to provide gender-neutral and female-gendered search results when one uses male-gendered search terms, just as systems making use of the semantic scoring of proximity would probably be better suited to provide job seekers with job listings that are *related* to their search queries but are not described using the same words.

The exact workings of the search algorithms of the Web sites cannot be fully known. For the purpose of this research we are interested in establishing whether employment Web sites in the Netherlands are causing harm.

A basic algorithm audit

The goal of this study is to evaluate whether the search algorithms of three particular employment Web sites counter and neutralize bias, or whether they reproduce and naturalize bias. With *neutralizing* we mean that when a gendered search term is entered, the job Web site retrieves job openings showing a gender-neutral job title made possible by working with stem or semantic scores. The exact matching of gendered-search terms to vacancies can reproduce and thereby normalize pre-existing societal bias.

To determine the existence and extent of bias and potential discrimination on employment Web sites, we conduct a very basic algorithm audit (Sandvig, *et al.*, 2014). More concretely, we examined three Web sites, among the largest employment Web sites in the Netherlands. The Web sites under scrutiny help explore how the search algorithms of employment sites, through their dealing with gendered language, can cause harm. We have decided to keep the Web sites anonymous as our primary aim is to create awareness about the issue of harm and search engines more broadly. Their parent companies were, however, informed of our results by the Netherlands Institute for Human Rights. While we happily disclose the Web sites to fellow academics, to our knowledge one of the sites has since changed their search algorithm as a result of our findings.

The normative claim in this paper is that employment Web sites should provide equal opportunities to all job seekers by showing users all vacancies relevant to their search query. In other words, these Web sites need to mitigate potential allocative harm resulting from gender bias in society. By no means does this type of ‘intervention’ solve the structural problem of gender stereotyping. It may, however, positively contribute to countering and slowly eroding these stereotypes.

For the audit we tried to reverse engineer the search algorithms (Diakopoulos, 2014) by comparing inputs and outputs. Specifically, we wanted to know how the search algorithms of these Web sites handle gendered job titles. By gendered job titles, we mean job titles that clearly target either male or female individuals, such as ‘secretaresse’, the Dutch title for a female administrative assistant. We compiled a list of gendered job titles and their neutral counterpart(s). These job titles were entered as queries on the three employment Web sites. Using these job titles, we searched for job vacancies with the most ‘relevancy’, which is the default mode of searching and therefore what likely was most commonly used.

The researcher performing the searches was not logged in to the given Web site and used *incognito* mode. Importantly, none of the Web sites asked job seekers to report their gender. In total, we used 15 clearly gendered titles and their neutral counterparts as queries on the three Web sites, making 30 queries in total. We chose to work with one gendered term (either male or female) and a neutral counterpart making up a pair, because not all terms employ both a female and a male variant. For instance, in Dutch usage there is no term ‘police woman’ accompanying ‘policeman’ and ‘police officer’.

To select job titles we used three criteria. First, the job title had to be clearly gendered, describing either female or male employees. Second, the job title had to be currently in use in the sense that the term for the job described had to be the common, most prevalent term on the contemporary job market. Last, job titles had to be spread across different sectors. To determine the various prevailing sectors in the job market, we looked at information released annually by Statistics Netherlands (CBS) releases to describe the state of the Dutch job market. In their yearly reports, they categorize the statistics on jobs by sector. The job titles we selected were evenly spread across the sectors of healthcare, education, commerce, manufacturing, travel, entertainment, public service and administrative.

Gendered job title	Neutral counterpart	English translation
Actrice	Acteur	Actress/actor
Brandweerman	Brandwacht	Firefighter
Gastheer	Gastvrouw / Gastheer	Host
Gastvrouw	Gastvrouw / Gastheer	Host
Kassière	Kassamedewerker	Cashier
Leidster kinderopvang	Kinderopvangmedewerker	Childcare worker
Leraar	Docent	Teacher
Lerares	Docent	Teacher
Politieman	Politieagent	Policeman/police officer
Schoonmaakster	Schoonmaakmedewerker	Cleaner
Secretaresse	Secretariaatsmedewerker	Administrative assistant
Stewardess	Steward(ess)	Flight attendant
Timmerman	Houtbewerker	Carpenter
Verpleegster	Verpleegkundige	Nurse
Verpleger	Verpleegkundige	Nurse

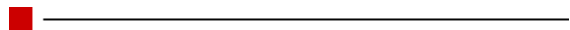
Table 1: The 15 chosen job titles plus their neutral counterparts. The seven titles in boldface were researched elaborately.

The selected job titles were then queried on the three job application Web sites. The search results of each query were manually copied and pasted into our MS Excel spreadsheet. For seven job titles (those in boldface in [Table 1](#)), we examined the job title and vacancy text of the first 10 search results (corresponding to the maximum number of results shown on the first page of one of the employment Web sites under scrutiny). These were chosen because they are clearly gendered and are very commonly used. For the rest of the job titles, we looked at the job title and the vacancy texts of the first three search results. Furthermore, a quantitative exploration was conducted by noting the total number of available job vacancies on the Web site for each of the 30 queries.

In examining the data, we paid attention to the scope of the search results shown for each query. When using a query that is clearly gendered to describe exclusively female or male employees, do the search results merely reflect this biased

gendered view of the position, or do they also include other, gender-neutral terms for the same position and thereby counter the pre-existing bias? We looked at the same thing when using gender-neutral queries: Do the search results echo the search query, or do they also include other terms describing the same position?

In examining the search results for each query, we looked at two things to determine whether they were gendered or gender-neutral: (a) the contents and wording of the job title; and (b) the vacancy texts. We could thus determine whether the search results merely echoed the (gendered) query used or whether they broadened the scope of the query by including gender-neutral vacancies. Second, in an effort to further substantiate our findings, we looked at the total number of vacancies available for a certain query on the Web site. By comparing the total number of vacancies for a gendered query and its gender-neutral counterpart, we gauged whether the algorithm treats the two queries as descriptors for the same job position or as two different things. In the event for the former, the total number of vacancies would be the same; for the latter, the number would be different.



The neutralization of gendered terms

In this section, we zoom in on the search results for some of our queries. First, we examined the three queries ‘leraar’, ‘lerares’ and ‘docent,’ respectively the male, female and neutral Dutch pronouns for ‘teacher’.

Percentage of neutralized search results for 'leraar' and 'lerares'

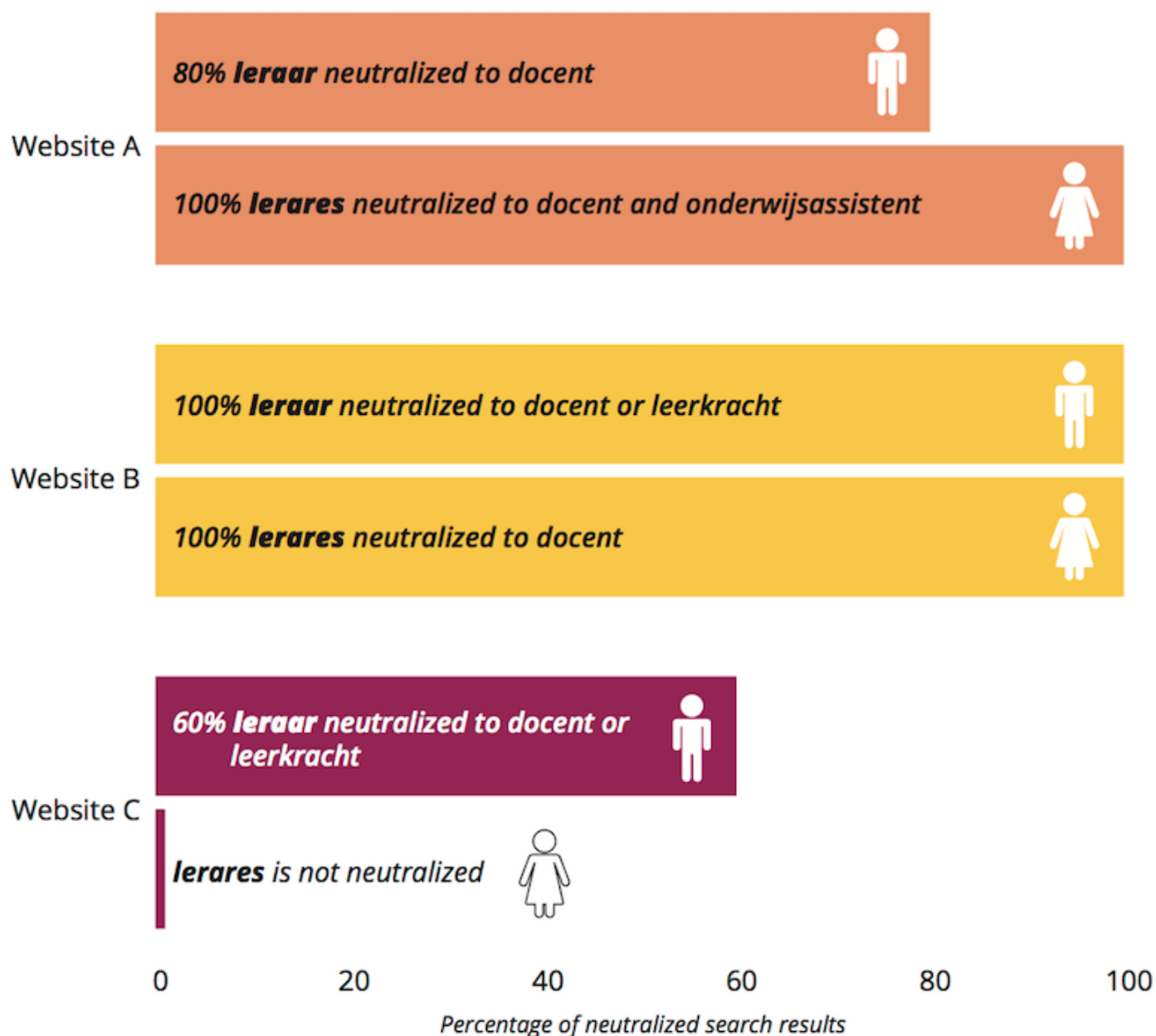


Figure 1: Percentage of neutrally worded search results when using gendered search terms. ‘Leraar’ is a male teacher and ‘lerares’ is a female teacher. The titles ‘docent’ and ‘leerkracht’ are also used to describe teachers and can mean both male and female teachers. The term ‘onderwijssistent’ is also gender neutral but means ‘teaching assistant’.

Figure 1 shows the percentage of naturally worded search results as an indication of whether the Web sites’ respective search algorithms neutralize the gendered terms ‘leraar’ and ‘lerares’. Here we qualitatively looked at the first ten search results (title and vacancy text) that the Web sites show when searching with the search terms indicated. The results indicate that employment Web sites A and B both neutralize a high percentage of the gendered search terms within the job vacancies we examined (80–100 percent). For example, when we look at Web site A, searching for ‘leraar’, 80 percent of all job vacancies are a gender-neutral vacancy, and the remaining 20 percent show a vacancy that echoes the gendered term ‘leraar’. For Web site B, all gendered search queries provide gender-neutral search results. This percentage

is much lower for Web site C, which also seems to reproduce pre-existing bias, neutralizing ‘leraar’ fairly often to search results including ‘docent’, whereas ‘lerares’ is never neutralized. In other words, searching for ‘lerares’, the Web site only shows employment opportunities that literally use the term ‘lerares’; the user is not exposed to vacancies for ‘docent’ or ‘leraar’.

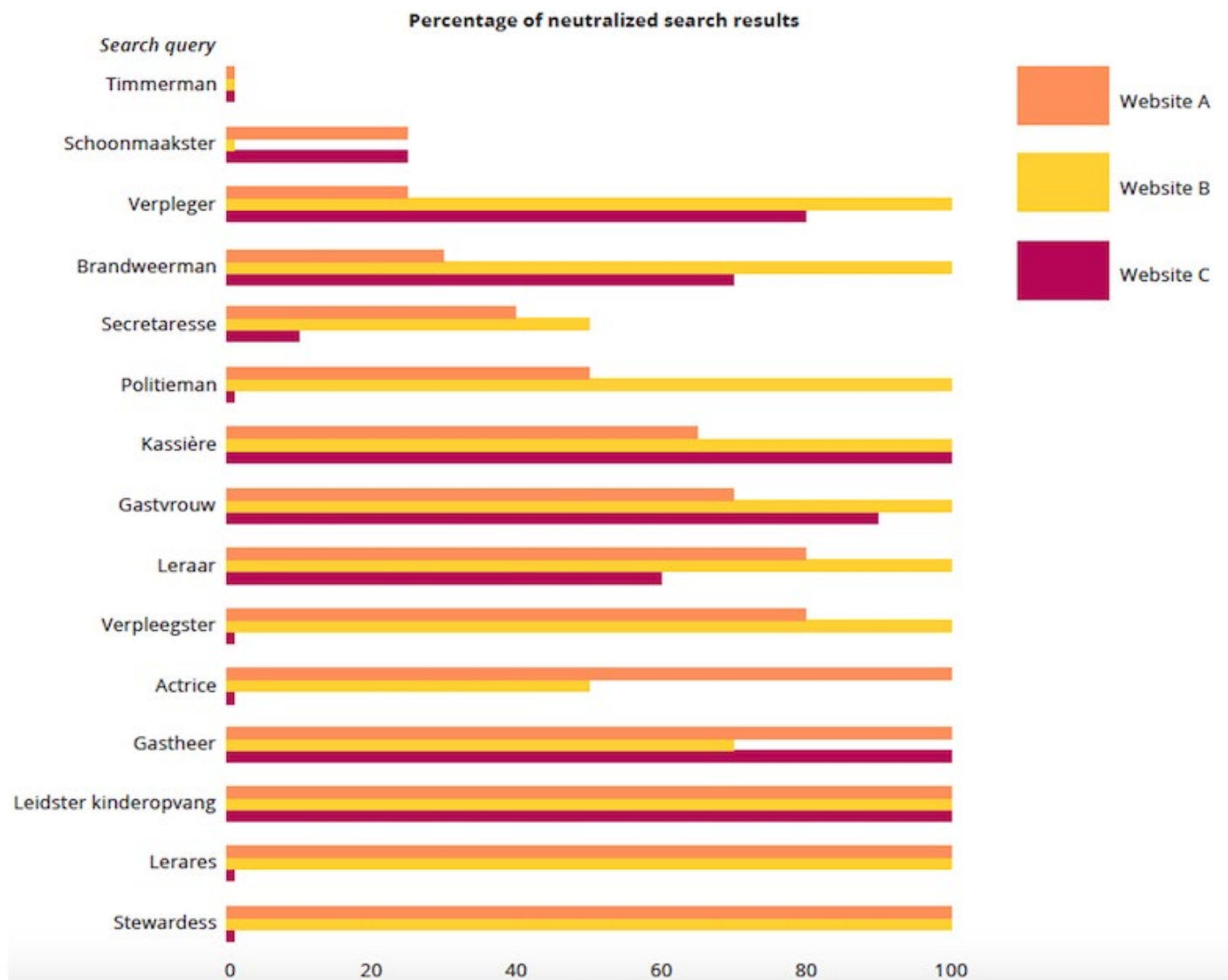


Figure 2: The y-axis shows the gendered search terms entered in the search bars of the Web sites. The x-axis shows the percentage of search results that are neutralized. Translation of search terms from top to bottom: carpenter (male), cleaner (female), nurse (male), firefighter (male), administrative assistant (female), policeman (male), cashier (female), host (female), teacher (male), nurse (female), actress (female), host (male), childcare worker (female), teacher (female) and flight attendant (female).

Zooming out, we took all fifteen gendered queries into account. [Figure 2](#) shows what percentage of the search results included a gender-neutral job title when searching with a gendered term. For instance, when you search for [verpleegster] (female Dutch noun for nurse) or [verpleger] (male Dutch noun for nurse) on Web site B, you see results only for the job title ‘verpleegkundige’ (neutral Dutch noun for nurse) in the top 10 search results. In line with our earlier finding, the neutralization percentages for Web site C are considerably lower than those of the other two Web sites. Web site B always neutralizes most of the gendered search terms.

Despite returning a high number of job listings, the search terms [*secretaresse*], [*schoonmaker*] and [*verpleegster*] are not neutralized very often on any of the employment Web sites. In all cases, the search term [*timmerman*] only retrieved vacancies that literally repeat ‘*timmerman*’.

A quantitative perspective

In an attempt to validate these findings, we also looked at the number of results retrieved for all gendered terms and their gender-neutral counterparts. Here, our hypothesis held that a Web site returning a similar number of vacancies for a gendered term and a gender neutral term probably employs a search algorithm in which both terms are included within a single category in the taxonomy. It ensures that when searching with the gendered term, all vacancies where the gender-neutral counterpart appears will be included in the results. It indicates whether people are presented with results beyond the exact search term. Presenting job seekers with a large pool of search results that exceed their gendered search would be desirable. In that event, the search algorithm plays a role in mitigating the job seeker’s own pre-existing biases. The results of this comparison can be found in Figures 3, 4 and 5.

Total number of search results for gendered terms ‘*leraar*’ and ‘*lerares*’ and neutral term ‘*docent*’

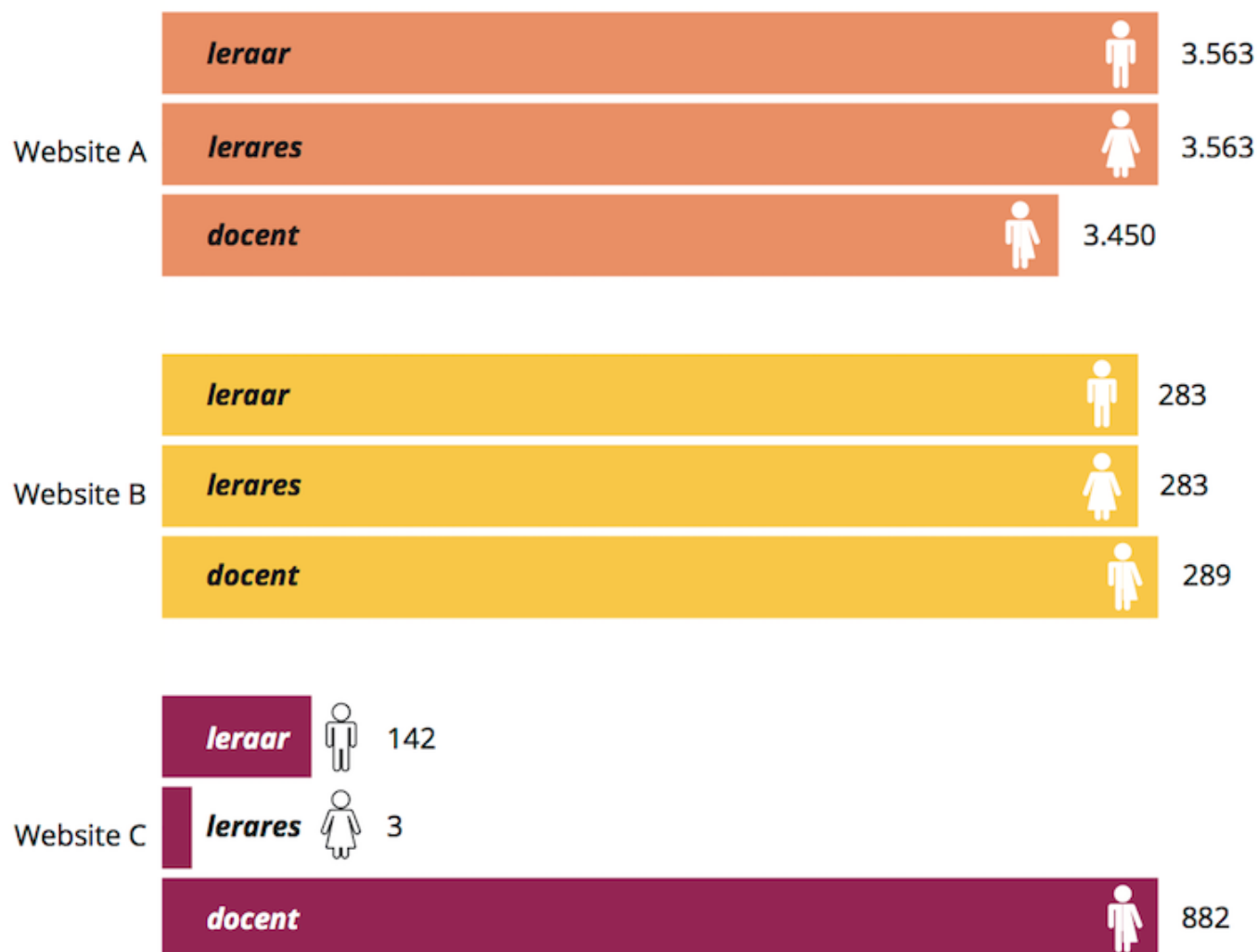


Figure 3: Comparison of number of vacancies retrieved for ‘leraar’ (male teacher), ‘lerares’ (female teacher) and ‘docent’ (neutral form of teacher).

Figure 3 shows that Web sites A and B return almost as many vacancies for the gendered terms ‘leraar’ and ‘lerares’ and their gender neutral counterpart ‘docent’. Web site C, however, shows striking differences in the number of retrieved vacancies. Here the gender neutral term ‘docent’ retrieves 882 results, whereas ‘lerares’ produces only three and ‘leraar’ 142. Another noteworthy finding is that Web sites A and B retrieve the exact same number of vacancies for ‘leraar’ as for ‘lerares’. This finding strongly suggests that the search algorithm pools these two search terms into the same job category.

Total number of search results for gendered terms ‘schoonmaakster’ and ‘schoonmaker’ and neutral term ‘schoonmaakmedewerker’

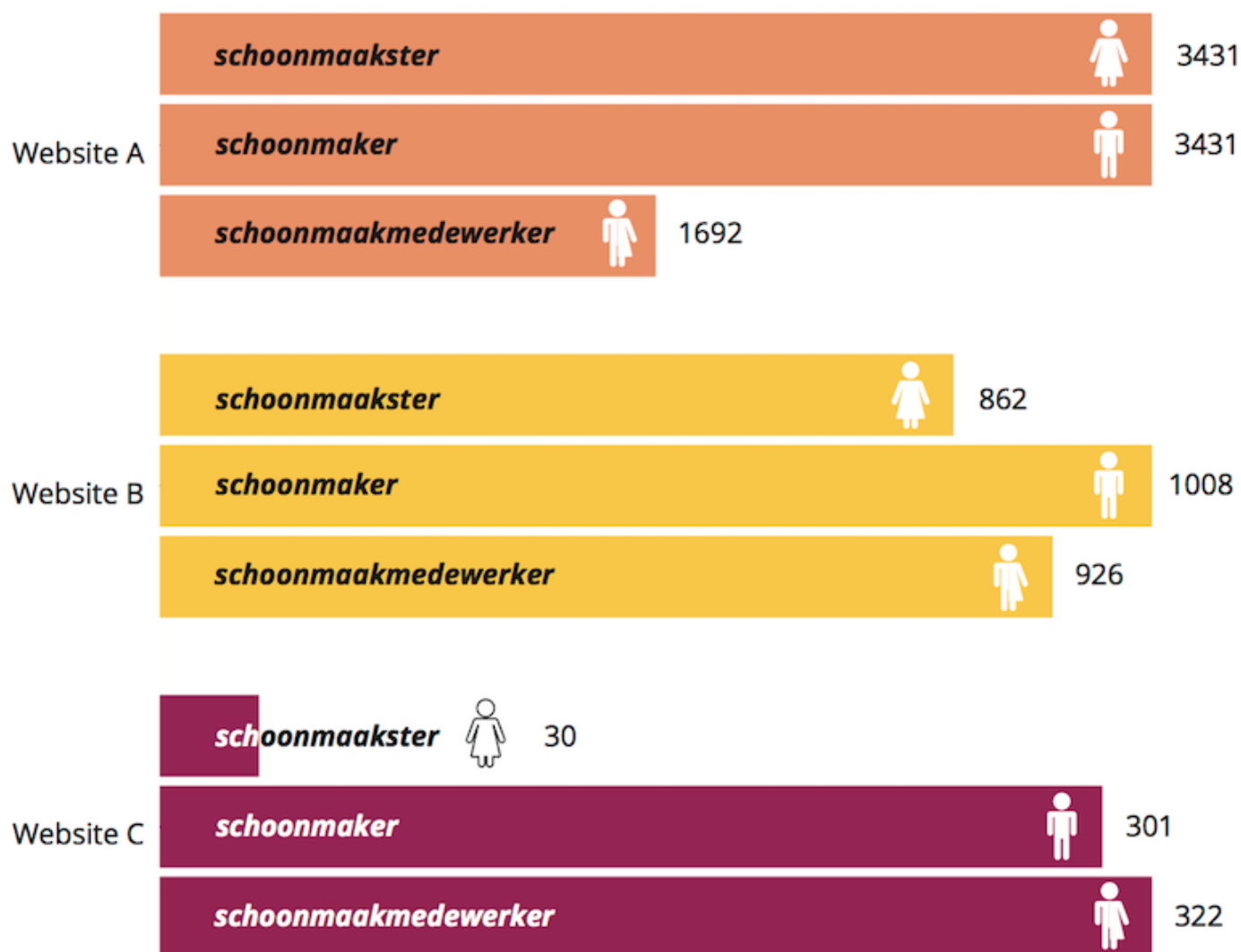


Figure 4: Comparison of the number of vacancies retrieved for ‘schoonmaakster’ (female cleaner), ‘schoonmaker’ (male cleaner) and ‘schoonmaakmedewerker’ (neutral form of cleaner).

Figure 4 compares the search terms [schoonmaakster], [schoonmaker] and [schoonmaakmedewerker]. Here we notice telling discrepancies on all Web sites. What one first notices is Web site A retrieving exactly the same number of vacancies for the female term 'schoonmaakster' as it does for the male term 'schoonmaker'. This strongly suggests that A's algorithm classifies the two terms as indicating the same category of job. Therefore, it has the potential to mitigate pre-existing bias. Further inspection of the retrieved vacancies revealed that Web sites A and B both neutralize 'schoonmaakster' to 'schoonmaker'. This yields more results than 'schoonmaakmedewerker', most likely indicating that this term has been naturalized to the 'neutral term'. The originally masculine term 'schoonmaker' to denote a cleaner is now seemingly understood as neutral. Again it is evident that Web site C neutralizes search terms the least.

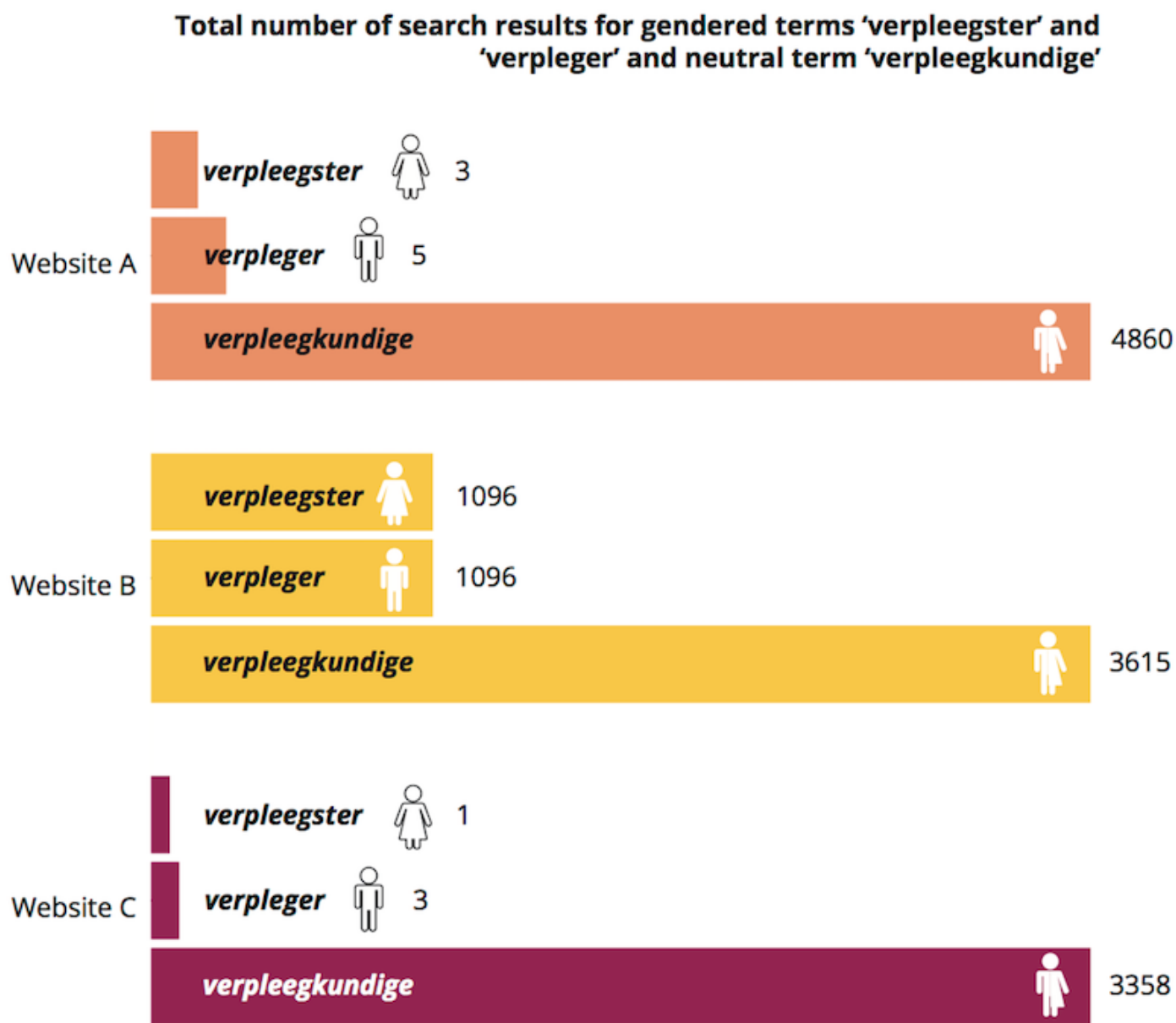


Figure 5: Comparison of the number of vacancies retrieved for 'verpleegster' (female nurse), 'verpleger' (male nurse) and 'verpleegkundige' (neutral form of nurse).

Figure 5 demonstrates that Web sites A and C retrieve a very small number of vacancies for the female and male search terms, whereas there is a high number of vacancies when searching with the neutral term. These results suggest that the

three different search terms are not divided into the same job category by the search algorithm. As in [Figure 3](#), Web site B retrieves the same number of vacancies for both the female and the male search term, suggesting that the algorithm classifies the two terms under the same job category. Overall, the neutral term ‘verpleegkundige’ is by far the dominant term for the job category of nurse.

These studies’ results suggest that Web site C uses a simple search algorithm that retrieves only literal matches for the entered query. By contrast, Web sites A and B each use a more complex search algorithm that is able to group gender-neutral and gendered job titles when describing a similar job function. As a result, these functions end up in the same pool even though they are referred to in different ways. This observation can plausibly explain why the number of job openings retrieved for the male and female search terms is exactly the same for Web site A in [Figures 3 and 4](#). This also applies to Web site B in [Figures 3 and 5](#). Their search algorithms probably work either by relating search terms to a pre-established taxonomy using word stems and word families or by assigning semantic scores to jobs based on “word embeddings” and thus retrieving similar results.

If one searches with a gender-neutral job title, none of the Web sites retrieve a gendered job title in their top ten search results. This suggests that gendered job titles are neutralized and, conversely, gender neutral job titles are not gendered. Here again the distribution of information on employment opportunities is negatively impacted. These online sites produce “distributive consequences” in relation to individuals’ exposure to employment opportunities and actively shape the labor market (Kim, 2020).

Ideally, a job seeker using gendered terms is exposed to a large pool of job vacancies that would extend beyond positions that echo a (gendered) search query on a particular Web site. Arguably, the Web site would do well to broaden the scope by presenting the user with *all* relevant employment opportunities (gendered *and* neutral). Whether job Web sites contribute to this effort — and if they do, to what extent — depends on how their search algorithm works.

Conclusion

Although much research has focused on complex algorithms and artificial engines in the job market for candidate selection and matching through curriculum vitae, there is a need to extend critical research to determine how search engines on employment Web sites do harm to their job-seeking users. This explorative study has taken a first step in this direction. It has examined the influence of the search algorithm used by employment Web sites for the Netherlands on the visibility of employment opportunities. Although we have witnessed the platformization of the Web (Helmond, 2015), we suspect that these types of Web sites, which still enjoy considerable reach, will not be displaced by social media platforms in the foreseeable future.

The results of our algorithm audit indicate that employment Web sites A and B use more complex search algorithms that probably work with a word-stem taxonomy or semantic proximity score. As such, they were seen to pool together search terms that can be classified as typically male or female to indicate the same job category. Furthermore, they returned gender neutral search results in response to a query containing a gendered search term. Such a response, however, did not happen the other way around: a gender-neutral search term did not prompt gendered search results.


A semantic link, or taxonomy, in the algorithm seems to work only in one direction, in the sense that it converts gendered search terms into gender-neutral search results but does not convert gender-neutral search terms into gendered search results. These algorithms, mitigating bias through searches with gendered search terms, may actively neutralize language and serve as a corrective for bias and gender stereotypes circulating in society.

In contrast, our results suggest that Web site C uses a fairly simple search algorithm with search results limited to those which literally match the used search term. Here, the bias of the job seeker is reaffirmed without an alternative to such ideas being provided. As such, Web site C contributes to the naturalization of gendered language. It reproduces gender stereotypes and prejudices regarding certain occupations and fails to make visible the entirety of relevant employment opportunities. A conversation with representatives of the Web site’s parent company that occurred after the commissioned report was published has confirmed our suspicions in this regard. They indicated they would be adding a taxonomy to the search engine to expand its search results.

Our research focused on the search algorithms of employment Web sites. We did not examine the auto-complete suggestions in the search bars, although these prompts should also be taken into consideration. LinkedIn, for example, changed its search algorithm to remove female-to-male prompts when people searching for female names were prompted

to arrive at similar sounding male names (Day, 2016). Moreover, the language of the vacancy texts themselves can also exclude certain groups from applying for such positions. Job Web sites in the Netherlands often do not take responsibility for the texts uploaded to their Web sites (Inspectie SZW, 2021).

In this paper we have focused on the technical operation of search algorithms to see whether they deliver all relevant job vacancies using both gendered and gender-neutral terms despite the use of either gendered or gender-neutral search terms. Here, we would stress that, in some cases in our corpus, female search terms were translated into ‘neutral’ terms that are actually male in origin. In contrast, male nouns were not always translated into gender-neutral nouns. We propose future research to explore this aspect further. For now, we surmise that, effectively, the search engine here erases female job titles such as ‘lerares’ and treats (some) male nouns as neutral. In doing so, it normalizes the position of men — whether unconsciously or not, whether unintentionally or on purpose.

To conclude, more attention needs to be paid to the role of tech intermediaries in distributing information about employment opportunities. As we have shown, even simple algorithms may cause harm. Specifically, employment Web sites should attend to how their search engines handle gender-inflected language so that job seekers will not be faced with a situation in which they are not seeing relevant employment opportunities. As we have demonstrated, a representational harm can be transformed into an allocative harm. Of course, these more inclusive search engines merely tackle the symptom rather than fixing the structural issue of gender stereotypes. An important step, however, that can be taken right now is to challenge these stereotypes and to incrementally move the needle with regard to the connection between specific genders and particular occupations. 

About the authors

Dr. **Karin van Es** is assistant professor of Media and Culture Studies at Utrecht University and lead researcher at Utrecht Data School. Her research explores the impact of datafication and algorithmization on culture and society.
E-mail: k[dot] f [dot] vanes [at] uu [dot] nl

Daniel Everts has a Master’s of Arts in the field of film and television studies. He is currently writing a thesis at Utrecht University on Instagram, the work of Bernard Stiegler and the subjects of mnemotechnology and human consciousness. His other interests lie in the critical examination of discriminatory practices and bias in Western society across interdisciplinary boundaries.
E-mail: d [dot] everts [at] uu [dot] nl

Iris Muis has a background in law and international relations and is currently project manager at Utrecht Data School. She specializes in data ethics and has developed multiple instruments for assessing the (ethical) impact of data practices.
E-mail: i [dot] m [dot] muis [at] uu [dot] nl

Acknowledgements

The authors would like to thank Marieke van Santen and Arthur Vankan for their contributions to the commissioned research project on which this paper in part draws. We also thank Didi van Zoeren for the design of the figures.

This paper draws on commissioned research that Utrecht Data School conducted for the Netherlands Institute for Human Rights (NIHR).

Notes

[1](#) Zuiderveen Borgesius, 2018, p. 10.

[2](#) *Ibid.*

[3](#) Lammerant, *et al.*, 2018; O’Neil, 2016, p. 24; Bowker and Star, 1999, pp. 46–47.

[4](#) Gitelman and Jackson, 2013, p. 3.

[5](#) Lammerant, *et al.*, 2018; O’Neil, 2016, p. 24; Bowker and Star, 1999, pp. 46–47.

6. O’Neil, 2016, p. 23.
7. D’Ignazio and Klein, 2020, p. 39.
8. Grommé, *et al.*, 2019, pp. 60–62.
9. d’Alessandro, *et al.*, 2017, p. 124.
10. Seaver, 2017, p. 5.
11. d’Alessandro, *et al.*, 2017, p. 124.
12. Dyer, 2002, p. 126.
13. Perez, 2019, pp. 6–7.
14. Foucault, 2007, p. 161; Hall, 1992, p. 291.
15. Sturken and Cartwright, 2009, p. 16.

References

- Ifeoma Ajunwa and Daniel Greene, 2019. “Platforms at work: Automated hiring platforms and other new intermediaries in the organization of work,” In: Steven P. Vallas and Anne Kovalainen (editors). *Work and labor in the digital age*. Bingley: Emerald, pp. 61–91.
doi: <https://doi.org/10.1108/S0277-283320190000033005>, accessed 15 July 2021.
- Miranda Bogen and Aaron Rieke, 2018. “Help wanted: An examination of hiring algorithms, equity, and bias,” *Upturn* (December), at <https://www.upturn.org/reports/2018/hiring-algorithms/>, accessed 3 May 2021.
- Geoffrey C. Bowker and Susan Leigh Star. 1999. *Sorting things out: Classification and its consequences*. Cambridge, Mass.: MIT Press.
doi: <https://doi.org/10.7551/mitpress/6352.003.0001>, accessed 15 July 2021.
- Armand Chatard, Serge Guimont and Delphine Martinot. 2005. “Impact de la féminisation lexicale des professions sur l’auto-efficacité des élèves: Une remise en cause de l’universalisme masculin?” *L’année psychologique*, volume 105, number 2, pp. 249–272, and at https://www.persee.fr/doc/psy_0003-5033_2005_num_105_2_29694, accessed 15 July 2021.
- Le Chen, Le, Ruijun Ma, Anikó Hannák and Christo Wilson, 2018. “Investigating the impact of gender on rank in resume search engines,” *CHI ’18: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, paper number 651, pp. 1–14.
doi: <https://doi.org/10.1145/3173574.3174225>, accessed 15 July 2021.
- Kate Crawford, 2017. “The trouble with bias,” at <https://blog.revolutionanalytics.com/2017/12/the-trouble-with-bias-by-kate-crawford.html>, accessed 3 May 2021.
- Brian d’Alessandro, Cathy O’Neil and Tom LaGatta. 2017. “Conscientious classification: A data scientist’s guide to discrimination-aware classification,” *Big Data*, volume 5, number 2, pp. 120–134.
doi: <https://doi.org/10.1089/big.2016.0048>, accessed 15 July 2021.
- Jeffrey Dastin, 2018. “Amazon scraps secret AI recruiting tool that showed bias against women,” *Reuters* (10 October), at <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G>, accessed 3 May 2021.
- Amit Datta, Michael Carl Tschantz and Anupam Datta, 2015. “Automated experiments on ad privacy settings: A tale of opacity, choice, and discrimination,” *arXiv:1408.6491* (17 March), at <https://arxiv.org/abs/1408.6491v2>, accessed 3 May 2021.
- Matt Day, 2016. “LinkedIn changes search algorithm to remove female-to-male name prompts,” *Seattle Times* (8

September), <https://www.seattletimes.com/business/microsoft/linkedin-changes-search-algorithm-to-remove-female-to-male-name-prompts/>, accessed 15 July 2021.

Nick Diakopoulos, 2014. “Algorithmic accountability reporting: On the investigation of black boxes,” *Tow Center for Digital Journalism, Columbia University* (10 July).
doi: <https://doi.org/10.7916/D8ZK5TW2>, accessed 15 July 2021.

Catherine D’Ignazio and Lauren F. Klein. 2020. *Data feminism*. Cambridge, Mass.: MIT Press.
doi: <https://doi.org/10.7551/mitpress/11805.001.0001>, accessed 15 July 2021.

Richard Dyer, 2002. *The matter of images: Essays on representation*. London: Routledge.

Magdalena Formanowicz, Sylwia Bedyska, Aleksandra Cislak, Friederike Braun and Sabine Sczesny, 2013. “Side effects of gender-fair language: How feminine job titles influence the evaluation of female applicants,” *European Journal of Social Psychology*, volume 43, number 1, pp. 62–71.
doi: <https://doi.org/10.1002/ejsp.1924>, accessed 15 July 2021.

Michel Foucault, 2007. “The meshes of power,” In: Jeremy W. Crampton and Stuart Elden (editors). *Space, knowledge and power: Foucault and geography*. Translated by Gerald Moore. Hampshire: Ashgate, pp. 153–162.
doi: <https://doi.org/10.4324/9781315610146>, accessed 15 July 2021.

Batya Friedman and Helen Nissenbaum. 1996. “Bias in computer systems,” *ACM Transactions on Information Systems*, volume 14, number 3, pp. 330–347.
doi: <https://doi.org/10.1145/230538.230561>, accessed 15 July 2021.

Tarleton Gillespie, 2014. “The relevance of algorithms,” In: Tarleton Gillespie, Pablo J. Boczkowski and Kirsten A. Foot (editors). *Media technologies: Essays on communication, materiality, and society*. Cambridge, Mass.: MIT Press.
doi: <https://doi.org/10.7551/mitpress/9780262525374.003.0009>, accessed 15 July 2021.

Lisa Gitelman and Virginia Jackson. 2013. “Introduction,” In: Lisa Gitelman (editor). *“Raw data” is an oxymoron*. Cambridge, Mass.: MIT Press, pp. 1–14.
doi: <https://doi.org/10.7551/mitpress/9302.003.0002>, accessed 15 July 2021.

Francisca Grommé, Sophie Emmert, Noortje Wiezer and Claartje Thijs. 2019. “Digitale arbeidsmarktdiscriminatie: Inzicht in de Risico’s Op Arbeidsmarktdiscriminatie Door de Inzet van Recruitment Technologieën in werving en selectie,” at <https://repository.tudelft.nl/islandora/object/uuid%3A4d78130c-5270-4c42-b321-6f1d8e380e00>, accessed 3 May 2021.

Stuart Hall, 1992. “The West and the rest: Discourse and power,” In: Stuart Hall and Bram Gieben (editors). *Formations of modernity*. Cambridge: Polity, pp. 275–332.

Anne Helmond, 2015. “The platformization of the Web: Making Web data platform ready,” *Social Media + Society* (30 September).
doi: <https://doi.org/10.1177/2056305115603080>, accessed 15 July 2021.

Inspectie SZW, 2021. “Eindrapportage discriminatiemogelijkheden online platforms,” *Publicatie, Ministerie van Sociale Zaken en Werkgelegenheid*, at <https://www.inspectieszw.nl/publicaties/publicaties/2021/01/22/eindrapportage-discriminatiemogelijkheden-online-platforms>, accessed 9 June 2021.

Pauline T. Kim, 2020. “Manipulating opportunity,” *Virginia Law Review*, volume 106, number 4, at <https://www.virginialawreview.org/articles/manipulating-opportunity/>, accessed 3 May 2021.

Hand Lammerant, P.H. Blok and Paul De Hert, 2018. “Big data besluitvormingsprocessen en sluiptwegen van discriminatie,” *Nederlands Tijdschrift Voor de Mensenrechten*, volume 43, number 1, pp. 3–24, and at https://cris.vub.be/ws/portalfiles/portal/43119057/pdh18_hlpbbig_data_sluiptwegen_NJCM_.pdf, accessed 15 July 2021.

Jack Morse, 2017. “Google Translate might have a gender problem,” *Mashable* (30 September), at <https://mashable.com/2017/11/30/google-translate-sexism/>, accessed 3 May 2021.

Safiya Umoja Noble, 2018. *Algorithms of oppression: How search engines reinforce racism*. New York: New York University Press.

Cathy O’Neil, 2016. *Weapons of math destruction: How big data increases inequality and threatens democracy*. New York: Crown.

Andi Peng, Besmira Nushi, Emre Kiciman, Kori Inkpen, Siddharth Suri and Ece Kamar, 2019. “What you see is what you get? The impact of representation criteria on human bias in hiring,” *Proceedings of the Seventh AAAI Conference on Human Computation and Crowdsourcing*, pp. 125–134, and at <https://ojs.aaai.org/index.php/HCOMP/article/view/5281>, accessed 3 May 2021.

Caroline Criado Perez, 2019. *Invisible women: Data bias in a world designed for men*. New York: Abrams Press.

Irene de Pous, 2020. “Vanaf Nu Een ‘Redactrice’,” *OnzeTaal.Nl*, at <https://onzetaal.nl/nieuws-en-dossiers/weblog/vanaf-nu-een-redactrice>, accessed 3 May 2021.

Manish Raghavan, Solon Barocas, Jon Kleinberg and Karen Levy, 2020. “Mitigating bias in algorithmic hiring: Evaluating claims and practices,” *FAT* ’20: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, pp. 469–481.
doi: <https://doi.org/10.1145/3351095.3372828>, accessed 15 July 2021.

David Sánchez and Antonio Moreno, 2004. “Automatic generation of taxonomies from the WWW,” In: Dimitris Karagiannis and Ulrich Reimer (editors). *Practical aspects of knowledge management. Lecture Notes in Computer Science*, volume 3336. Berlin: Springer, pp. 208–219.
doi: https://doi.org/10.1007/978-3-540-30545-3_20, accessed 15 July 2021.

Christian Sandvig, Kevin Hamilton, Karrie Karahalios and Cederic Langbort. 2014. “An algorithm audit,” In: Seeta Peña Gangadharan with Virginia Eubanks and Solon Barocas (editors). *Data and discrimination: Collected essays*. Washington, D.C.: New America Foundation, pp. 6–10, and at <http://www-personal.umich.edu/~csandvig/research/An%20Algorithm%20Audit.pdf>, accessed 15 July 2021.

Nick Seaver, 2017. “Algorithms as culture: Some tactics for the ethnography of algorithmic systems,” *Big Data & Society* (9 November).
doi: <https://doi.org/10.1177/2053951717738104>, accessed 15 July 2021.

Dagmar Stahlberg and Sabine Sczesny, 2001. “Effekte des generischen Maskulinums und alternativer Sprachformen auf den gedanklichen Einbezug von Frauen,” *Psychologische Rundschau*, volume 52, number 3, pp. 131–140.
doi: <https://doi.org/10.1026/0033-3042.52.3.131>, accessed 15 July 2021.

Jane G. Stout and Nilanjana Dasgupta, 2011. “When he doesn’t mean you: Gender-exclusive language as ostracism,” *Personality and Social Psychology Bulletin*, volume 37, number 6, pp. 757–769.
doi: <https://doi.org/10.1177/0146167211406434>, accessed 15 July 2021.

Marita Sturken and Lisa Cartwright, 2009. *Practices of looking: An introduction to visual culture*. Second edition. Oxford: Oxford University Press.

Arianal Tobin, 2019. “Employers used Facebook to keep women and older workers from seeing job ads. The federal government thinks that’s illegal,” *ProPublica* (24 September), at <https://www.propublica.org/article/employers-used-facebook-to-keep-women-and-older-workers-from-seeing-job-ads-the-federal-government-thinks-thats-illegal>, accessed 3 May 2021.

Dries Vervecken, 2013. “The impact of gender fair language use on children’s gendered occupational beliefs and listeners’ perceptions of speakers,” doctoral thesis, Freien Universität Berlin, at https://refubium.fu-berlin.de/bitstream/handle/fub188/5473/Dries_Vervecken_Ph_D_The_Impact_of_Gender_Fair_Language_Use__For_UB.pdf?sequence=1, accessed 3 May 2021.

Sandra Wachter, 2020. “Affinity profiling and discrimination by association in online behavioural advertising,” *Berkeley Technology Law Journal*, volume 35, number 2, pp. 367–430, and at https://btlj.org/data/articles2020/35_2/01-Wachter_WEB_03-25-21.pdf, accessed 3 May 2021.

Maranke Wieringa, 2021. Correspondence with Karin van Es (3 June).

Frederik Zuiderveen Borgesius, 2018. “Discrimination, artificial intelligence, and algorithmic decision-making,” *Council of Europe*, at <https://rm.coe.int/discrimination-artificial-intelligence-and-algorithmic-decision-making/1680925d73>,

accessed 15 July 2021.

Editorial history

Received 4 May 2021; revised 10 June 2021; accepted 11 July 2021.



This paper is licensed under a [Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License](https://creativecommons.org/licenses/by-nc-nd/4.0/).

Gendered language and employment Web sites: How search algorithms can cause allocative harm
by Karin van Es, Daniel Everts, and Iris Muis.

First Monday, volume 26, number 8 (August 2021).

doi: <http://dx.doi.org/10.5210/fm.v26i8.11717>