The Relationship Between AI Recruitment and Gender Bias: A Literature Review

Sofie Rietdijk (2071797)

Personeelwetenschappen, Tilburg University

Bachelor Thesis

Supervisor: M. Spanouli

Second reader: R. Bosman

Bachelor Thesis

May 31, 2024

Abstract

This systematic literature review provides a comprehensive overview of the relationship between AI recruitment and gender bias. This study aims to analyze the extent to which the use of AI in the recruitment process maintains, strengthens, reduces, or eliminates biases against gender. Fifteen articles were found systematically using the databases WorldCat and GoogleScholar. Inclusion criteria, such as a timeframe from 2018 up to and including 2024 and peer-reviewed English-written articles only were used. Based on these criteria, 15 articles were found and analyzed. The literature presented a twofold perspective regarding the relationship between AI recruitment and gender bias. One side showed a negative relationship and the other side a positive relationship. One reason for the positive relation that was given was the ability of AI to remove gender identifications in the recruitment process. Most of the literature supported the negative relationship. The most common reason for the negative relationship was the input of biased data. Since most of the literature supported the negative relationship, it might be interesting to investigate possible interventions and strategies in the future, to overcome this negative relationship. The lack of focus on a specific country or region, the pace of technological advancements and the lack of intersectionality, could be possible limitations of this study. Therefore, another recommendation for future research could be to take these aspects in consideration. To overcome the negative relationship between AI recruitment and gender bias organizations could implement information sessions and trainings on how to be unbiased.

Keywords: *Artificial Intelligence, recruitment, gender, bias, algorithms*

The Relationship Between AI recruitment and Gender Bias

Transformative technologies, like big data and artificial intelligence (AI), were introduced by the fourth revolution (Zhang & Chen, 2023). According to Budhwar et al. (2022), AI-based technologies in human resource management (HRM) are transforming the organization of work, decision-making and problem-solving. A significant area within HRM where the influence of AI is evident lies within recruitment (Chen, 2023). AI implementation in recruitment can potentially improve a company's competitiveness, since it can provide a more thorough understanding of talent compared to competitors (Chen, 2023). This is because AI can notice and find both potential silent job candidates and exceptional job candidates (Black & Van Esch, 2020).

What the relationship between AI recruitment and gender bias is, is highlighted from a twofold perspective. On one side, Raveendra et al. (2020) mention that AI provides optimism for removing human biases and discrimination which are prevalent in the hiring process. One of the driving factors behind AI recruitment is the capability to automate allegedly objective forms of assessment. AI hiring firms, firms that use AI recruitment, strive to offer employers a method to quickly shift through vast numbers of applicants without bias. These firms imply that their AI recruitment can eliminate considerations of gender and race from the hiring process (Drage & Mackereth, 2022). In this literature review, the focus will be on bias related to gender. Gender bias can be defined as: "The difference in ratings or perceptions of men and women with identical qualifications." (Isaac et al., 2009, p.1). Research found that gender bias appears within the hiring process. Rouwenhorst and Trisko (2005) found that gender biases in hiring decisions are primarily evident in the final decision-making stage, where both men and women predominantly choose male candidates for hiring.

Given the fact that gender bias is present within traditional recruitment processes, it is essential to investigate how AI might affect this bias. The other side of the perspective shows that decisions made by AI in the hiring process are shaped by the initial data it receives. If the foundational data is unfair, the resulting algorithms may sustain bias or discrimination and thus foster the potential for inequality (Chen, 2023). According to Chen (2023), this undermines the concept of social justice, which results in moral and economic damage to those affected by bias or discrimination. This moral and economic damage will lead to decreased economic efficiency, which results in decreased production of services and goods (Chen, 2023). Whether AI recruitment indeed relates to gender bias or not, and if so, in what form, is a particularly relevant subject to study.

The relationship between AI recruitment and gender bias has theoretical and practical relevance. This literature review will be an extension of the existing literature that is to be found about this topic. This is because it gives a structured overview of literature, about the relationship between AI recruitment and gender bias investigated within the timeframe of 2018 until the present. The existing literature doesn't yet provide an overview like this within this time frame. The rapid advancement of AI technologies in recent years necessitates an up-to-date examination of their implications in recruitment processes, and therefore, such a study is needed. This literature review will fill in this gap and therefore be additional to existing literature.

Apart from the theoretical relevance, this literature review also provides practical relevance for organizations and HR professionals. Organizations can benefit greatly from understanding how AI recruitment relates to gender bias to ensure fair hiring practices. Since fair hiring practices enhance organizational performance and effectiveness (King, 2005) and unfair hiring practices can ultimately lead to a reduction of economic efficiency (Chen, 2023), understanding this relationship is of great importance for organizations. HR professionals play an essential role in the design and implementation of recruitment strategies. This literature review has the potential to provide HR professionals with a comprehensive understanding of the potential challenges and opportunities that come along with the implementation of AI in recruitment processes. With this knowledge, HR professionals can make well-informed decisions, implement best practices and advocate for fair and ethical recruitment practices within their organizations.

Since a more profound understanding of the relationship between AI recruitment and gender bias brings both theoretical and practical relevance, it is relevant to review the selected literature, after applying the inclusion and exclusion criteria, to investigate this relationship. The following research question is formulated for the search: "*How does AI recruitment relate to gender bias?*" The answer to this question can be formulated after the literature has been searched, selected, and analyzed.

In the following section, the construct definitions of 'AI recruitment' and 'gender bias' will be elaborated on, along with additional background on the variables and the scope of the review. After that, the methods, results, discussion, and the conclusion that can be drawn from this literature review will follow.

Theoretical Framework

AI recruitment and gender bias are concepts defined in many ways in the existing literature. In this section, the definitions of how these concepts are interpreted in this literature review will be given. Also, two critical theoretical perspectives are discussed that can help understand the relationship between AI recruitment and gender bias.

AI Recruitment

Artificial intelligence can be defined as a system's ability to interpret data correctly, learn from it and use that learning in tasks through flexible adoption (Haenlein & Kaplan, 2019). AI recruitment means that this intelligence is used within recruitment processes. It uses algorithms that can evaluate a resume based on qualifications, experience and skills and provide candidates with positive or negative feedback in response. It analyzes the applicant's characteristics to generate a profile that predicts whether an applicant possesses the required skills and expertise that the company is looking for (Brishti & Javed, 2020). Within AI recruitment, AI influences each stage of the recruitment process. According to Fritts and Cabrera (2021), the influences in each stage are as follows:

1. *Sourcing:* AI can help to construct targeted searches to generate an initial pool of potential candidates.
2. *Screening:* AI can rapidly assess and rank resumes or predict the compatibility of a candidate's personality within the organizational culture.
3. *Interview:* AI can automate the initial interview round and contribute to the evaluation of candidates.
4. *Selection:* AI can impact the final selection process by recommending the top candidates from the remaining pool.

Recruitment professionals claim that because of the use of AI in their recruitment, they can find more qualified candidates and can reduce the time spent sorting through resumes (Fritts & Cabrera, 2021). Advantages like these, stem from AI's ability to process vast amounts of information and make decisions at speeds far beyond human capacity (Black & Van Esch, 2020).

Certain downsides of AI recruitment are addressed in the literature as well. The implementing of AI within the recruitment process, requires significant time (Son & Oh, 2023). The literature also addresses another downside, namely gender bias. Gender bias could occur because of AI recruitment when the data used to teach algorithms is biased against a particular gender (Njoto et al., 2022).

Gender Bias

Gender bias is characterized by the favoritism towards one gender over another. (Mhadgut, 2023). Friedman and Nissenbaum (1996) use the term bias to refer to behavior that systematically discriminates against specific individuals or groups in favor of others. Gender bias in recruitment encompasses any form of unequal treatment or differential consideration exhibited towards male and female applicants (Arceo-Gómez & Campos-Vázquez, 2022). When gender bias occurs in the recruitment process in an organization, it can lead to several negative consequences. According to King (2005), research indicates that teams characterized by diversity tend to outperform homogenous groups in innovation, problem-solving, and decision-making. Therefore, addressing gender bias in recruitment is crucial not only for promoting fairness but also for enhancing organizational performance and effectiveness.

The concept of gender bias within the field of AI falls under the umbrella term algorithmic bias. Algorithms detect patterns within large datasets, deriving rules for decision-making and automated predictions (Hall & Ellis, 2023). They serve as quantifiers, meaning they can make hiring decisions or recommendations by quantifying values which are fed into the system by the programmer (Fritts & Cabrera, 2021). Algorithmic bias can be defined as unfair discrimination of groups such as gender and race by algorithms (Lee, 2022). A theory that investigates unfair discrimination based on gender, is the gendered organizational theory (Clark-Saboda & Lemke, 2023).

Gendered Organizational Theory

According to Rodríguez and Guenther (2022), the gendered organizational theory can be defined as a theory that pays attention to how organizations foster and tackle emerging and evolving gender(ed) inequalities. It is an understanding of organizations as environments where gender dynamics and gender order are (re)created and brings the gender lens to discussion about organizations (Rodríguez & Guenther, 2022). Key ideas of the gendered organizational theory center around gender as a social construct that underpins inequalities in working life (Rodríguez & Guenther, 2022). Researchers have used the gendered organizational theory to identify, evaluate, and eliminate (invisible) gender discrimination and inequity in organizations (Clark-Saboda & Lemke, 2023). With the gendered organizational theory in mind, potential gender bias deriving from AI recruitment could, therefore, be identified and eliminated. This theory helps identify the fact that AI recruitment is not just a technical process but one embedded in social and organizational contexts where gender biases are formed and perpetuated. The theory prompts us to consider how existing gender norms and possible inequalities within organizations may be reinforced by AI systems, if these biases are not addressed.

Social-Cognitive Theory

The social-cognitive theory, as described by Bandura (1986), is a theory that helps us understand how people behave by looking at three dimensions that influence each other: what people think and feel (cognitive and affective factors), what they do (behavioral patterns) and what's happening around them (environmental factors) These three things are connected in a way that they all affect each other. People evoke varied responses from their social environment based on characteristics like age, gender, size and physical attractiveness. Similarly, they elicit diverse reactions based on their socially assigned roles and status. These social responses influence how individuals perceive themselves and others, reinforcing or diminishing environmental bias (Bandura, 1999). Since the algorithms in AI recruitment are shaped by the initial data it receives (Chen, 2023), this data must be unbiased. Subjective human involvement comes into selecting features and models used in the algorithms and how they include identifiable factors such as race and gender (Veale & Binns, 2017).

So, by recognizing that individuals elicit varied responses from their social environment based on characteristics such as their gender, the social-cognitive theory emphasizes how these biases can be perpetuated or mitigated through AI algorithms. This is because AI algorithms are shaped by the initial data they receive, including subjective human involvement. The social-cognitive theory highlights the importance of designing AI recruitment systems that prioritize unbiased data and decision-making processes, ultimately contributing to reducing gender bias in recruitment practices.

In this literature review, the focus will be on literature that has investigated the relationship between AI recruitment and gender bias. This will include research from psychology, sociology, computer science and business administration.

Methods

To answer the research question, a systematic literature review on the relationship between AI recruitment and gender bias has been conducted. A literature review distils the body of existing literature within a specific subject area and provides a comprehensive summary of the current state of knowledge within that field (Rowley & Slack, 2004). According to Rowley and Slack (2004), the systematic part of a literature review refers to the fact that knowledge about the topic is searched in a systematic and reproducible way. More specifically, a Critically Appraised Topic (CAT) is followed. A CAT is a summary that is structured in a standardized manner and contains research evidence organized around a research question. This summary aims to provide a critique on the research and a statement of the relevance of the results (Sadigh et al., 2012). This section will discuss which methodology is used to conduct this systematic literature review. It will provide a thorough and detailed overview of how, where and when the literature is found, and which inclusion and exclusion criteria were used.

The Literature Search

The databases WorldCat and Google Scholar were used for the literature search to find studies about the topic. The search in these databases was conducted within the time frame of

March 2024 up to and including May 2024. Specific search terms were picked to capture the broadest spectrum of literature related to the relationship between AI recruitment and gender bias. The search terms were combined in the following way: “AI recruitment AND gender bias” and “Artificial Intelligence recruitment AND gender bias”. Table 1 shows that Google Scholar gave n=206 results using the combination “AI recruitment AND “gender bias” and WorldCat n=1400 results. Google Scholar gave n=51 results using the combination “Artificial Intelligence recruitment” AND “gender bias” and WorldCat gave n=1100 results.

Table 1

Search Terms and Results per Database

Search terms	**Google Scholar**	**WorldCat**
“AI recruitment” AND “gender bias”	206	1400
“Artificial Intelligence recruitment” AND “gender bias”	51	1100

Included and Excluded Resources

Using the search terms in Table 1, the databases Google Scholar and WorldCat presented 2757 sources. The usable literature needed to be peer-reviewed and English written articles. Based on these criteria, the initial set of resources could be excluded from both databases combined (n=1985 excluded). This literature review only includes research that is investigated within the time span from 2018 until 2024. The studies may not be older than 2018 because AI recruitment emerged in 2018 (Fritts & Cabrera, 2021). Therefore, the resources that didn’t include research investigated within this timespan were therefore excluded (n=275 excluded). Since there is no specific focus on country or region, resources weren’t excluded based on this. After these inclusion criteria, 497 resources were left, which could be assessed on their title. Assessing the titles means that the title should give the impression that the article entails both the constructs of AI recruitment and gender bias. This is because this literature review investigates the relationship between these two constructs and not one of them on its own (n=444 excluded). The 53 resources that were left could be assessed based on their abstract. Based on their abstract, several resources could be excluded since the abstract clearly showed that this resource wasn't investigating the relationship between AI recruitment and gender bias (n=38 excluded). After this exclusion, 15 resources were left, all included for a full-text review. After conducting a full text review one resource could be excluded since it didn't provide an answer to the research question (n=1 excluded).

After all these criteria, 14 resources were identified through a database search and included in this literature review. During the full-text review of the article from Njoto (2020), one usable resource could be found during the reference check of this article, namely the one written by Costa et al. (2022). After including this last resource, 15 resources were analyzed and discussed for this literature review. The analysis and discussion of these 15 resources ultimately provided the necessary information to answer the research question. A systematic approach was used to analyze and integrate the findings from the different resources. This involved the identification of similarities, differences, and recurring patterns. All together were synthesized into meaningful insights about the relationship between AI recruitment and gender bias.

Figure 1

Search Scheme and Exclusion Criteria

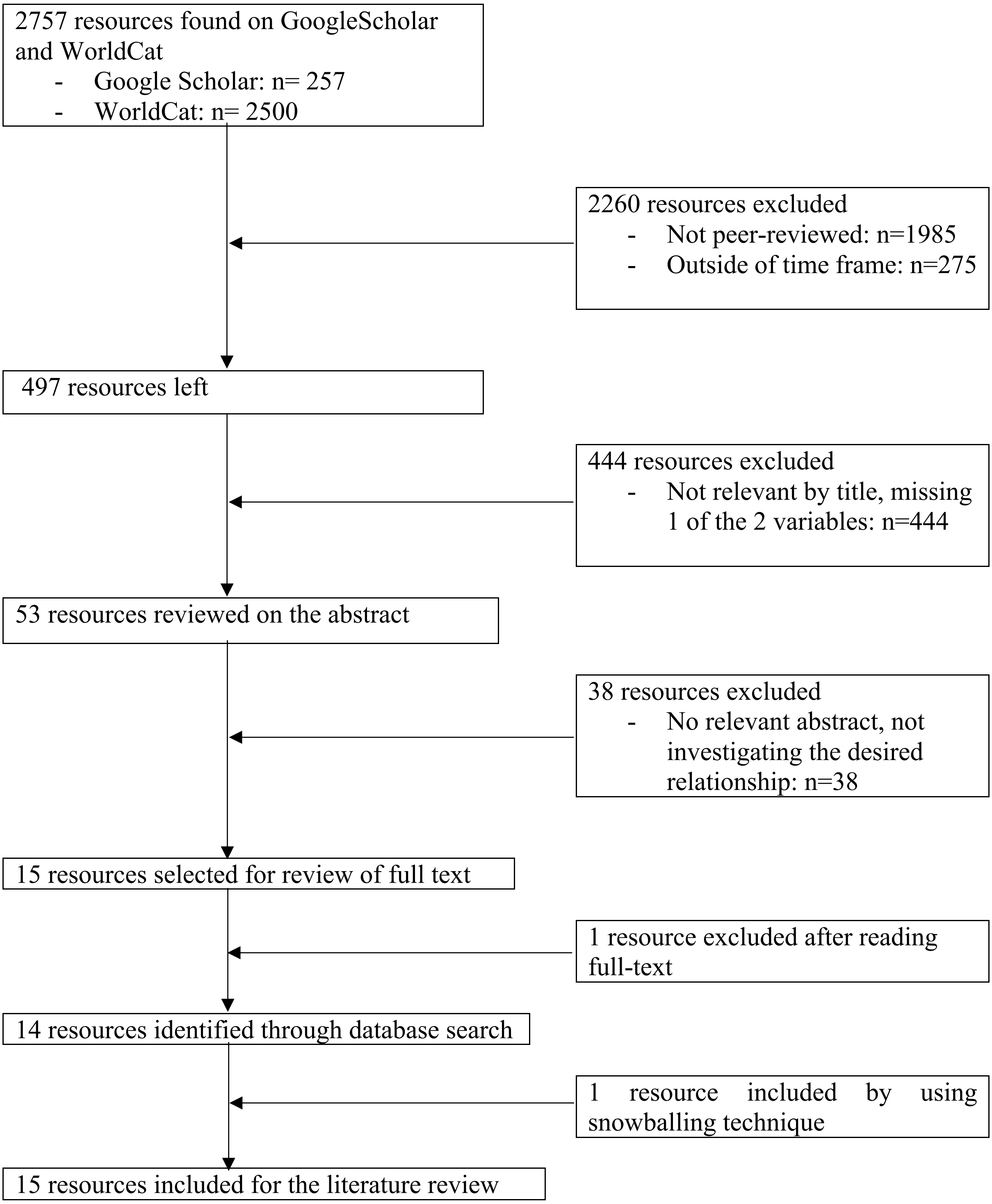
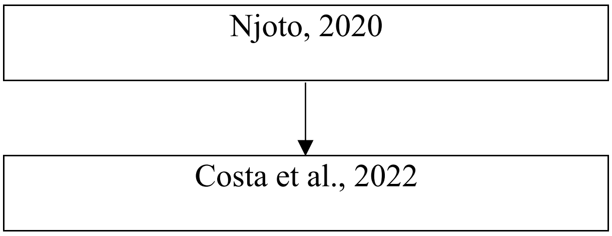


Figure 2

Resources Found Using the Snowballing Technique



Results

This section will provide an extensive analysis of the ultimate set of literature, after applying the inclusion and exclusion criteria for this literature review. It will clarify evident patterns, disparities and trends which underscore the relationship between AI recruitment and gender bias.

An Optimistic Approach

With the emergence of AI recruitment in 2018, the entire recruitment process of companies was revitalized (Fritts & Cabrera, 2021). According to Fritts & Cabrera (2021) several organizational benefits were discovered when implementing the intelligence of machines within the recruitment process. Such benefits are the saving of costs and time within the recruitment process and the ability to find more qualified candidates (Fritts & Cabrera, 2021). Furthermore, an optimistic approach emerged about AI's effect on avoiding or overcoming human bias within the recruitment process (Köchling & Wehner, 2020). Different studies were conducted on this topic and found that AI recruitment could, indeed, remove biases towards gender within the recruitment process (Hemalatha et al., 2021; Houser, 2019; Uma et al.2023).

Removal of Gender Identifications

Houser (2019) discusses in her paper that AI successfully reduces gender bias. This is explained by using different AI recruitment software as examples. She discusses that these software applications can remove names and gender identifications from resumes and can also remove faces and names from applicants' LinkedIn profiles, for instance. The removal of names and gender identifications leads to hiring more women (Houser, 2019). According to Houser (2019), organizations tend to choose men more often to come over for an interview even though they have the same credentials as women. When these gender identifications are hidden, the possibility of discrimination based on gender is removed (Houser, 2019).

Objective Evaluation

Building upon this optimistic approach towards the relationship between AI recruitment and gender bias, Hemalatha et al. (2021) and Uma et al. (2023) provide further insights into the positive impact of AI technologies in the recruitment process. Hemalatha and colleagues (2021) critically analyze the effect of AI on recruitment and selection in organizations. Their findings indicated that the use of AI in recruitment and selection has several positive organizational outcomes, including the removal of bias in the form of gender and ethnicity. Uma et al. (2023) conducted a scoping review to examine the existing literature. They found that the algorithms used for evaluation in resume screening ensure objectivity and therefore remove bias, among others, in the form of gender.

A Pessimistic Approach

After highlighting the potential benefits of AI in mitigating gender bias within the recruitment process, it is crucial to acknowledge that the literature also highlights another perspective on how the two concepts relate. Hemalatha et al. (2021), Houser (2019), and Uma et al. (2023) discuss the ability of AI to remove gender identifications and its objective nature as reasons for how AI can prevent, reduce, or even remove gender bias within the recruitment process. However, despite these perceived benefits, literature also highlights that AI-driven recruitment methods may inadvertently create, prolong, or compound gender bias.

A Misunderstanding of the Concept of Gender

Completely contradictory to the findings mentioned earlier (Hemalatha et al., 2021; Houser, 2019; Uma et al., 2023) are the findings of Drage and Mackereth (2022). They conducted research in which they examined different claims from AI hiring companies. One of those claims was that AI could objectively evaluate candidates by removing gender. They found that AI recruitment contributes to and sustains gendered biases. The reason for this is that efforts to remove gender from the recruitment system fail to grasp the true nature of these concepts. The AI hiring companies perceive gender as an isolated characteristic rather than acknowledging it as an integral component of broader systems of societal power and inequality. Consequently, they fail to address the systemic factors contributing to gender bias.

Biased Input Data

While Uma et al. 2023 claim that algorithms ensure objectivity, the findings of several studies completely contradict this suggestion (Albassam, 2023; Chen, 2023; Costa et al., 2022; Gupta & Mishra, 2023; Köchling & Wehner, 2020; Miasato & Silva, 2020; Njoto, 2020; Njoto et al., 2022; Ntoutsi et al., 2020; Peña et al., 2020). In each of these studies, a clear pattern becomes apparent regarding the role of AI recruitment in perpetuating gender biases. This pattern revolves around the input of biased data into the algorithm of the AI recruitment software.

Köchling and Wehner (2020) conducted a systematic literature review to provide a reliable picture of the existing knowledge about the fairness and discrimination potentials when using algorithmic decision-making in recruitment. Their main finding is that algorithmic decision-making within recruitment is not a universal solution for eradicating biases, although this is the reason why most companies implement it. Algorithms are vulnerable to biases in terms of gender if the algorithm is built upon inaccurate, biased, or unrepresentative input data. This is because algorithms replicate biases from the input data. They discuss that because of this, the process should be transparent so that both employees and candidates have the possibility to understand what is happening. Peña et al. (2020) support the evidence about biased training data with their experiment. The study was conducted using FairCVtest, an experimental framework for AI-based recruitment. The framework includes the creation of a dataset called FairCVdb, which consists of synthetic profiles with information typically found in job applicants' resumes. These profiles are scored with gender biases to simulate discrimination in the models targeted for candidate scoring in AI hiring processes. This framework revealed that standard AI algorithms exhibit gender bias when trained on biased data. Even though gender attributes were not explicitly provided, the algorithms can infer and exploit this information from other sources, such as face images. They discuss that this is not limited to specific cases but can also arise due to feature selection or unbalanced data representation, particularly when datasets lack diversity.

Albassam (2023) conducted a comprehensive analytical review of the current AI-based recruitment strategies by investigating industry reports and academic research. The review was conducted to critically evaluate the potential pros and cons of using AI in the recruitment process. They focus on the capability of AI to screen resumes and identify the best candidates. They discuss that algorithms, which can screen resumes, are trained on a large dataset of resumes, which is the basis for the identification of patterns and the prediction of the suitability of a candidate. Organizations using a biased dataset against gender in their AI recruitment creates algorithms which will perpetuate the same biases when selecting candidates. Costa et al. (2022) strengthen the conclusion of their research that AI recruitment is biased against gender by highlighting a practical example of the phenomenon mentioned by Albassam (2023). They highlight that Amazon created an algorithm to assess resumes and infer the best application using their workforce as training data. However, since most of their workforce were males, the algorithm exhibited systematic bias against female applicants. The algorithms assigned lower scores when encountering words such as ‘female’ and were more likely to reject their job applications.

Algorithmic Engineers

Building on this perspective is the literature review from Chen (2023), which was focused on AI recruitment and, more specifically, on algorithmic discrimination. His main conclusion was that bias in recruitment algorithms is evident in gender. The reason behind this is that algorithms stem from historical data. The personal preferences of algorithm engineers play a role in generating bias within these algorithms. If the engineers are biased, the algorithm is too (Chen, 2023).

Gupta and Mishra (2022) conducted a literature review on the ethical concerns of using AI in recruitment. They explain the reason behind the conclusion given by Chen (2023) on how the preferences of algorithm engineers play a significant role in generating gender bias within the algorithms of AI recruitment. They argue that only 22% of all AI professionals are women. Consequently, gender biases against women are perpetuated in AI technologies due to their underrepresentation in this sector (Gupta & Mishra, 2023). To link this to the conclusion of Chen (2023), most of the people working on algorithms are men, leading to an algorithm which is fed with bias against women. Miasato and Silva (2020) fully agree with the conclusions regarding the preferences of algorithmic engineers (Chen, 2023; Gupta & Mishra, 2023) in their article, since they claim that the use of AI in recruitment has opened a door for discrimination against gender, ethnicity, and orientation. They argue that this is because the programming is done by a person, and that when this person has prejudices this will reflect the machine’s decisions (Miasato & Silva, 2020).

Njoto (2020) further builds upon this perspective with the findings of her ‘James and Jane study’ findings. She conducted an experiment in which she created two identical resumes; the only aspect that differed was the name on these resumes. One resume was from ‘Jane’ and the other from ‘James’. She submitted these resumes to 40 companies using algorithmic recruiting. The findings of the experiment were as follows: Jane received 16 rejection emails, whereas James only received 12, and James received 15% more positive responses than Jane. Findings from such an experiment can be explained by research conducted by Ntoutsi et al. (2020). They explain how AI recruitment is prone to gender bias due to biased algorithms. According to Ntoutsi et al. (2020), the reason behind this is that algorithm developers might struggle to formulate assumptions objectively or rely on inaccurate selection criteria. So, when algorithms incorporate specific genders, they may inadvertently reinforce biased outcomes by correlating these attributes with the target variable. This phenomenon, therefore, underscores the potential for generating gender-biased algorithmic outcomes (Ntoutsi et al., 2020).

Gendered Language

Njoto and colleagues further investigated the relationship between AI recruitment and gender bias by conducting a sociological and data science-based case study, which showed that unconscious gender bias in human decision-making infiltrates algorithmic recruitment models (Njoto et al., 2022). They highlight that this phenomenon is exacerbated by ranking keywords and experiences based on gendered language in applications. How gendered language can lead to gender bias within AI recruitment is explained by Andrews and Bucher (2022). AI algorithms can also notice the differences in language between men and women resulting from societal norms. An example that Andrews and Bucher (2022) give is that women use collaborative language such as ‘we’ when describing projects, while men use ‘I’ more often when describing achievements. Consequently, when an algorithm is trained mainly on men, it will be biased to choose applicants using ‘I’ language on their resume rather than ‘we’ (Andrews & Bucher, 2022).

Table 1

Results

Author (s) & Year	**Keywords**	**Main findings**	**Implications**
Houser, 2019	AI, diversity, decision making, gender discrimination, unconscious bias	AI in recruitment is successful in reducing gender bias. It can eliminate gender identifications which will lead to the elimination of the possibility of discrimination based on gender.	More awareness of the risks involved in using data which do not represent a society at large and the need for continuous testing to mitigate biased results.
Hemalatha et al., 2021	Artificial Intelligence, recruitment, selection, automation, human resource management practices	The use of AI in recruitment has several organizational benefits, including the removal of bias in the form of gender.	AI has the potential to transform and improve the recruitment process, leading to more effective and efficient hiring practices.
Uma et al., 2023	Artificial intelligence, recruitment, chatbots, virtual interviews, resume screening	Algorithms used for evaluation in resume screening ensure objectivity and thus remove bias in the form of gender.	Need for organizations to adapt technological advancements and leverage AI tools effectively. Organizations need to consider factors such as generational preferences and cultural differences in their recruitment.
Köchling and Wehner, 2020	Algorithmic decision making in HRM, fairness, discrimination, perceived fairness, ethics	Algorithmic decision-making within recruitment is not a universal solution for eradicating biases, which is why most companies implement it. Algorithms are vulnerable to biases in terms of gender if the algorithm is built upon inaccurate, biased or unrepresentative input data.	Need for more sophisticated and theoretically grounded research in algorithm decision-making in HRM. But also, the necessity for ongoing debates and research on fairness and potential discrimination of algorithms
Chen, 2023	Algorithmic bias, AI-enabled recruitment, recruitment quality, discriminatory hiring practices	Recruitment algorithms' bias is evident in gender. The reason behind this is that algorithms stem from historical data. The personal references of algorithm engineers play a role in generating bias within these algorithms.	Need for a balanced approach that leverages the benefits of AI in recruitment while addressing and mitigating the ethical and discriminatory challenges associated with algorithmic decision-making processes.

Peña et al., 2020	Multimodel AI, discrimination, algorithms, fairness, automatic recruitment	Standard AI algorithms exhibit gender bias when trained on biased data. Even though gender attributes were not explicitly provided, the algorithms can infer and exploit this information from other sources, such as face images.	Need for transparency and accountability in automated recruitment systems to ensure that decisions are fair and unbiased. But also, discrimination-aware learning since these techniques can help mitigate biases by removing sensitive information from the decision making process.
Gupta and Mishra, 2022	Artificial Intelligence, recruitment, applications, ethical issues	Only 22% of all AI professionals are women. Consequently, gender biases against women are perpetuated in AI technologies due to their underrepresentation in this sector, and this is why AI recruitment consists of gender bias against women.	Need for ethical considerations to ensure fair and unbiased recruitment processes enabled by AI.
Miasato and Silva, 2020	Biased algorithms, discrimination in labor relations, artificial intelligence	The use of AI in recruitment has opened a door for discrimination against gender, ethnicity, and orientation. The reason for this is that a person does the programming, and when this person has prejudices, this will reflect the machine's decisions.	Need for education on ethical considerations in technology development. Individuals, particularly students in technology-related fields, should be prepared to handle the challenges posed by biased algorithms and understand the social impacts of technology.
Njoto, 2020	Artificial intelligence, algorithms, gender equality, hiring discrimination	Companies using algorithmic recruiting favored the resume with the name of a man more, than with the name of a woman on it, even though the rest of the resumes were identical. This shows that this algorithmic recruiting was biased against women.	Need for public policy responses and the re-examination of existing legal frameworks to address algorithmic biases and discrimination in hiring practices.
Njoto et al., 2022	Algorithmic bias, gender, recruitment, CV	Unconscious gender bias in human decision-making infiltrates in algorithmic recruitment models. This phenomenon is exacerbated by ranking keywords and experiences based on gendered language in applications.	Development of comprehensive training modules to help human resources personnel and hiring panels understand the complicated nature of bias as a first step to raising awareness and mitigating some of the bias.

Andrews and Bucher, 2022	Algorithm, gender discrimination, bias,	Gendered language can lead to gender bias within AI recruitment. The differences in language between men and women resulting from societal norms can also be noticed by AI algorithms. When an algorithm is trained mainly on men, it will be biased to choose applicants using more 'men language'.	Careful consideration of AI and algorithmic technologies in the hiring process is needed to ensure fairness and equal opportunities for all individuals.
Albassam, 2023	HRM technology, artificial intelligence, recruitment, AI-based recruitment strategies,	Organizations using a biased dataset against gender in their AI recruitment will create algorithms that perpetuate the same biases when selecting candidates.	Need for HR professionals to develop new knowledge and skills to effectively leverage AI in recruitment.
Costa et al., 2020	Machine learning, fairness, transparency, hiring, algorithmic justice	AI recruitment is biased against gender. Suppose the input data is focused on the existing workforce of an organization, which mostly consists of men. In that case, the AI algorithms will assign lower scores when words like 'female' are encountered in the job application.	Need for greater transparency in hiring processes to mitigate the negative impact of hiring algorithms.
Ntoutsi et al, 2020	Artificial intelligence, fairness, fairness-aware machine learning	AI recruitment is prone to gender bias due to biased algorithms. This is because algorithm developers might struggle to formulate assumptions objectively or rely on inaccurate selection criteria. So, when algorithms incorporate specific genders, they may inadvertently reinforce biased outcomes by correlating these attributes with the target variable.	Need to transcend conventional AI algorithms and incorporate ethical and legal principles into the entire lifecycle of AI systems, from design and training to deployment. This approach ensures social good while still benefiting from the potential of AI.

Drage and Mackereth, 2022	Artificial intelligence, recruitment, hiring, bias	AI recruitment contributes to and sustains gendered biases. The reason for this is that efforts to remove gender from the recruitment system fail to grasp the true nature of these concepts. The AI hiring companies perceive gender as an isolated characteristic rather than acknowledging it as an integral component of broader systems of societal power and inequality.	Need for a more critical and nuanced approach to the development and deployment of AI-powered HR tools, particularly about issues of bias, diversity, and inclusion.

Discussion

In this systematic literature review, specifically, this Critically Appraised Topic (CAT), the concepts of AI recruitment and gender bias were analyzed. This analysis was done based on 15 articles investigating the relationship between those two concepts. Therefore, this study aimed to provide an answer to the following research question: *"How does AI recruitment relate to gender bias?"* In other words, this literature review was conducted to determine to what extent the use of AI in the recruitment process would be able to maintain, strengthen, reduce, or even remove biases against gender.

As mentioned before, to answer this research question, 15 articles investigating this relationship were analyzed. After this analysis, it became clear that there is a distinct division within this literature. One side of the literature addresses the positive relationship between AI recruitment and gender bias, and the other side the negative relationship. Starting with the positive side. The literature discusses that the use of AI within the recruitment process can mitigate bias since it can remove gender identifications like profile pictures and names (Houser, 2019) and that it can ensure objectivity (Hemalatha et al., 2021; Uma et al., 2023). By this reasoning, it seems as if AI recruitment can mitigate or even remove gender bias within the recruitment process. For this reason, the two concepts can be seen as positively related.

Although this perspective seems promising, most of the literature contradicts the earlier findings of a positive relationship. Two main reasons are: biased input data and a misunderstanding of the concept of gender (Albassam, 2023; Chen, 2023; Costa et al., 2022; Gupta & Mishra, 2023; Köchling & Wehner, 2020; Miasato & Silva, 2020; Njoto, 2020; Njoto et al., 2022; Ntoutsi et al., 2020; Peña et al., 2020). This literature says that when the algorithm of an AI recruitment software is fed with biased input data, it will perpetuate gender biases. Njoto et al. (2022) and Andrews and Bucher (2022) suggest that this phenomenon is exacerbated by ranking keywords and experiences in an application based on gendered language. So, the claim that AI recruitment can ensure objectivity is undermined by algorithms operating based on the input data, which is not always unbiased and objective.

The study of Drage and Mackereth (2022) challenges the assertion made by Houser (2019) that AI can remove gender identifications and, therefore, mitigate gender bias. Drage and Mackereth (2022) found that when removing gender identifications from applications, the concept of gender is often misunderstood, since it is perceived as an isolated characteristic rather than an integral component of broader systems of societal power and inequality. Consequently, the systematic factors contributing to gender bias cannot be addressed.

It can be concluded that the studies finding a positive relationship between AI recruitment and gender bias, have even more studies on the other side finding the exact opposite results. Since most studies within this literature review found a negative relationship between AI recruitment and gender bias, the positive relationship should be further investigated. Further investigation could lead to more clarity on the contradictory results or possibly eliminate them. The two folded perspective indicates the need for a more nuanced understanding of the interplay between AI technologies and societal biases, suggesting further research to address the underlying factors perpetuating gender bias in recruitment processes.

Limitations

This literature review was designed and conducted to gather a body of literature, ultimately aiming to provide an answer to the research question. Despite this, some limitations can be addressed in this research. The first limitation is that no resources were excluded based on their country or region. The problem that could have arisen with not excluding resources based on this criterion is that the relationship between AI recruitment and gender bias may differ across different cultural and institutional contexts. For example, what is considered as gender bias in one country may not be perceived the same way in another. This is because different cultures have distinct attitudes towards gender roles. This limitation could, therefore, affect the generalizability of the findings of this literature review to diverse populations. Another limitation of this literature review could be that the intersectional nature of gender bias in AI recruitment is overlooked. This means that not all the studies consider how other factors intersect with gender to compound or mitigate biases. For example, an algorithm may exhibit different biases based on gender and race. The result of this limitation is that the studies may provide an incomplete or limited understanding of the complexities of gender bias in AI recruitment. The third limitation is the pace of technological advancements. AI technology evolves at an unprecedented pace, with new algorithms and techniques being developed continuously. Since the most recent study investigated in this literature review is from 2023, AI recruitment tools might have significantly improved by that time. These advancements could mitigate or eliminate some of the biases identified in this literature review, making the results unreliable.

Future Research Recommendations

From the results of this literature review, it became clear that the literature presents a twofold perspective concerning the relationship between AI recruitment and gender bias. One side of the perspective argues how these two concepts are positively related, and the other side shows how these two are negatively related. Most of the literature supported the side that argued the negative relationship. But is there a possible way to overcome this negative relationship? To investigate this, an interesting field for future research could be research into the possible interventions and strategies to mitigate gender bias within AI recruitment. For instance, such an intervention or strategy could be to give an awareness training to AI developers, since biased input data is the most common reason for gender bias in AI recruitment according to the investigated literature. This research could consist of an experiment in which AI developers need to make a pre- and post-training assessment to measure the changes in awareness after such training was implemented.

Another interesting topic for future research could be how to create more transparency and accountability in algorithms used in the recruitment process. By examining how to ensure transparency and accountability, biases perpetuating gender inequalities can be identified and addressed. An example of this is to focus research on the development or evaluation of mechanisms to enhance this transparency and accountability, such as 'algorithmic auditing' practices. These involve conducting systematic reviews of AI systems to assess their performance and identify potential biases. Another recommendation for future research could be to investigate how gender bias in AI recruitment varies across different industries and functions. Various functions within industries have varying requirements and expectations. It might be interesting to explore how AI systems assess candidates. For example, technical roles may face different biases compared to administrative roles since these roles often require skills and experiences that may be gender stereotyped. Investigating this could entail quantitative analysis of large datasets of job postings, applicants' profiles, and hiring outcomes across various industries and job functions.

Practical and Theoretical Implications

This literature review gives an up-to-date examination of the implications of the rapid advancements of AI technologies in the recent years. The twofold perspective found in the investigated literature, showing positive and negative relationships between AI recruitment and gender bias, suggests that theoretical models need to take the conditions under which AI can either mitigate or exacerbate bias into account. This indicates a need for theories that specify the conditions that influence these outcomes.

Organizations can gain valuable insights from this study. Since most of the literature cites biased input data as the primary reason for the negative relationship between AI recruitment and gender bias, the input data should require the most attention. Tackling problems like biased input data could possibly lead to a reduction of gender bias within AI recruitment. Information sessions and training on how to be unbiased could be a first step to tackle the problem at the core. Organizations could implement different interventions to prevent their algorithms from being biased, by ensuring that the input data is unbiased. Two interviews were conducted with HR-professionals (*see Appendix A and B)*. When this practical implication was presented to them, and they were asked if they could think of possible interventions. Interviewee 1 mentioned that an awareness training on how to be unbiased could be a possible intervention (see Appendix A). Interviewee 2 also came up with such an intervention and explicitly mentioned that a training like that should be given to the senior staff, since they are less educated on topics like gender diversity (see Appendix B). Raising more awareness on how to be unbiased among employees whose work it is to feed the algorithm with input data, could be a possible solution to avoid the negative relationship between AI recruitment and gender bias.

Conclusion

This systematic literature review investigated the relationship between AI recruitment and gender bias. The study was conducted to answer the following research question: *"How does AI recruitment relate to gender bias?"*. The answer is twofold. The most significant part of the literature showed how AI recruitment and gender bias are negatively related, with biased input data as the most given reason. The literature also gave reasons for why the two concepts were positively related, namely the fact that AI could ensure objectivity and remove gender identifications from applications. To tackle this negative relationship at the core, organizations could consider information sessions and training on how to be unbiased for their employees.

References

Albassam, W. A. (2023). The power of artificial intelligence in recruitment: an analytical review of current AI-based recruitment strategies. *International Journal Of Professional Business Review*, *8*(6), 1-25. https://doi.org/10.26668/businessreview/2023.v8i6.2089

Andrews, L & Bucher, H. (2022). Automating discrimination: AI hiring practices and gender inequality. *Cardozo law review, 44*(1), 145-202. https://cardozolawreview.com/automating-discrimination-ai-hiring-practices-and-gender-inequality/

Arceo-Gómez, E. O., & Campos-Vázquez, R. M. (2022). Gender bias in evaluation processes. *Economics Of Education Review*, *89*(1), Article e102272 https://doi.org/10.1016/j.econedurev.2022.102272

Bandura, A. (1999). A social cognitive theory of personality: Handbook of personality (2nd ed.,pp. 154-196). New York: Guilford Publications.

Black, J. S., & Van Esch, P. (2020). AI-enabled recruiting: What is it and how should a manager use it? *Business Horizons*, *63*(2), 215–226. https://doi.org/10.1016/j.bushor.2019.12.001

Budhwar, P., Malik, A., Silva, M., & Thevisuthan, P. (2022). Artificial intelligence challenges and opportunities for international HRM: a review and research agenda. *The International Journal of Human Resource Management*, 33(3), 1065 - 1097. https://doi.org/10.1080/09585192.2022.2035161.

Haenlein, M., & Kaplan, A. (2019). A brief history of artificial intelligence: On the past, present, and future of artificial intelligence. *California Management Review*, *61*(4), 5-14. https://doi.org/10.1177/0008125619864925

Chen, Z. (2023). Ethics and discrimination in artificial intelligence-enabled recruitment practices. *Humanities And Social Sciences Communications*, *10*(1), Article e567 https://doi.org/10.1057/s41599-023-02079-x

Clark-Saboda, J. L., & Lemke, M. (2023). Revisiting Acker's gendered organizational theory: what women overcome to stay in the superintendency. *Frontiers in Education*, *8*, Article e1160225 https://doi.org/10.3389/feduc.2023.1160225

Costa, A., Cheung, C. and Langenkamp, M. (2020). Hiring fairly in the age of algorithms (Publication No. 340662575) [Research Paper, Cornell University]. https://doi.org/10.48550/arXiv.2004.07132

Drage, E., & Mackereth, K. (2022). Does AI debias recruitment? Race, gender, and AI's "eradication of difference". *Philosophy & Technology*, *35*(4), 35-89. https://doi.org/10.1007/s13347-022-00543-1

Friedman, B., & Nissenbaum, H. (1996). Bias in computer systems. *ACM Transactions on Information Systems*, *14*(3), 330–347. https://doi.org/10.1145/230538.230561

Fritts, M., & Cabrera, F. (2021). AI recruitment algorithms and the dehumanization problem. *Ethics And Information Technology*, *23*(4), 791–801. https://doi.org/10.1007/s10676-021-09615-w

Gupta, A., & Mishra, M. (2022). Ethical concerns while using artificial intelligence in recruitment of employees. *Business Ethics And Leadership*, *6*(2), 6–11. https://doi.org/10.21272/bel.6(2).6-11.2022

Hall, P., & Ellis, D. J. (2023). A systematic review of socio-technical gender bias in AI algorithms. *Online Information Review*, *47*(7), 1264–1279. https://doi.org/10.1108/oir-08-2021-0452

Hemalatha, A., Kumari, P. B., Nawaz, N., & Gajenderan, V. (2021). Impact of artificial intelligence on recruitment and selection of information technology companies.

Proceedings of The International Conference on Artificial Intelligence and Smart Systems, *India*, 60-66. https://doi.org/10.1109/ICAIS50930.2021.9396036

Houser, K. (2019). Can AI solve the diversity problem in the tech industry? Mitigating noise and bias in employment decision-making. *Social Science Research Network*, *22*(1), 290-354.

Isaac, C., Lee, B., & Carnes, M. (2009). Interventions that affect gender bias in hiring: A Systematic review. *Academic Medicine*, *84*(10), 1440–1446. https://doi.org/10.1097/acm.0b013e3181b6ba00

King, J. (2005). Benefits of women in science. *Science*, *308*(5722), 601. https://doi.org/10.1126/science.1112550

Köchling, A., & Wehner, M. C. (2020). Discriminated by an algorithm: a systematic review of discrimination and fairness by algorithmic decision-making in the context of HR recruitment and HR development. *Business Research*, *13*(3), 795–848. https://doi.org/10.1007/s40685-020-00134-w

Lee, J. (2022). Algorithmic bias and the New Chicago School. *Law, Innovation and Technology*, *14*(1), 95–112. https://doi.org/10.1080/17579961.2022.2047520

Mhadgut, N. S. C. (2023). Gender Bias and Constitution of India. *International Journal Of Advanced Research in Science, Communication And Technology*, *3*(1), 85–86. https://doi.org/10.48175/ijarsct-9266

Miasato, A., & Silva, F. R. (2020). Artificial intelligence as an instrument of discrimination in workforce recruitment. *Acta Universitatis Sapientiae. Legal Studies, 8*(2), 191–212. https://doi.org/10.47745/ausleg.2019.8.2.04

Njoto, S. (2020). Gendered bots? Bias in the use of artificial intelligence in recruitment. Technical Report. The Policy Lab, The University of Melburne.

Njoto, S., Cheong, M., Lederman, R., McLoughney, A. J., Ruppanner, L., & Wirth, A. (2022). Gender bias in AI recruitment systems: A sociological-and data science-based case study. *IEEE International Symposium On Technology And Society (ISTAS) 1,* 1-7. https://doi.org/10.1109/istas55053.2022.10227106

Ntoutsi, E., Fafalios, P., Gadiraju, U., Iosifidis, V., Nejdl, W., Vidal, M., Ruggieri, S., Turini, F., Papadopoulos, S., Krasanakis, E., Kompatsiaris, I., Kinder-Kurlanda, K., Wagner, C., Karimi, F., Fernández, M., Alani, H., Berendt, B., Kruegel, T., Heinze, C., . . . Staab, S. (2020). Bias in data-driven artificial intelligence systems: An introductory survey. *Wiley Interdisciplinary Reviews. Data Mining And Knowledge Discovery,10*(3), Article 1356 https://doi.org/10.1002/widm.1356

Peña, A., Serna, I., Morales, A., & Fierrez, J. (2020) Bias in multimodal AI: Testbed for fair automatic recruitment. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, *USA,* 28-29. https://doi.org/10.1109/cvprw50498.2020.00022

Raveendra, P. V., Satish, Y. M., & Singh, P. (2020). Changing landscape of recruitment industry: A study on the impact of artificial intelligence on eliminating hiring bias from recruitment and selection process. *Journal Of Computational And Theoretical Nanoscience*, *17*(9), 4404–4407. https://doi.org/10.1166/jctn.2020.9086

Rodríguez, J. K., & Guenther, E. (2022). Gendered organization theory. *Oxford Research Encyclopedias*. https://doi.org/10.1093/acrefore/9780190224851.013.53

Rowley, J., & Slack, F. (2004). Conducting a literature review. *Management Research News*, *27*(6), 31–39. https://doi.org/10.1108/01409170410784185

Son, S., & Oh, J. (2023). A study on the performances of AI recruitment system: A case study on domestic and abroad companies. *Hanguk Jillo Changeop Gyeongyeong Yeongu*, *7*(2), 137–155. https://doi.org/10.48206/kceba.2023.7.2.137

Uma, V. R., Velchamy, I., & Upadhyay, D. (2023). Recruitment analytics: Hiring in the era of artificial intelligence. *Emerald Publishing Limited eBooks* (pp. 155–174). https://doi.org/10.1108/978-1-80382-027-920231008

Veale, M., & Binns, R. (2017). Fairer machine learning in the real world: discrimination without collecting sensitive data. *Big Data & Society*, *4*(2), Article e20539517. https://doi.org/10.1177/2053951717743530

Zhang, J., & Chen, Z. (2023). Exploring human resource management digital transformation in the digital age. *Journal Of The Knowledge Economy*, *14*(2), 1065-1086. https://doi.org/10.1007/s13132-023-01214-y

Appendix A

Summary Interview Interviewee 1

The first interview was held with an HR-advisor at the Ministry of Internal Affairs. As an HR-advisor she is the primary point of contact of the management for all their questions regarding personnel. At the Ministry of Internal Affairs AI recruitment is used in the form of CV screening and talent sourcing. By talent sourcing is meant that certain algorithms help recruiters to identify online profiles, on LinkedIn for instance, to find the best candidates and contact them. CV screening is used for the traineeship positions because thousands of applications are received for only 15 traineeships per year. The interviewee was surprised by the fact that the biggest part of the literature suggests that AI recruitment and gender bias are negatively related to each other. She believes that biases are due to the environment people grew up in and was convinced that algorithms could always exclude those biases. The Ministry of Internal Affairs wants to prevent themselves from any form of bias and promote diversity and have therefore appointed a diversity group called ‘BZK breed’. This group provides information sessions and programs. Everyone who works within recruitment is obligated to follow a training in which you learn to be unbiased in hiring. These biases entail cultural and gender biases. The interviewee also mentions that she thinks that these kind of information sessions and trainings, in which you learn to be unbiased, are the best intervention to overcome the negative relationship between AI recruitment and gender bias. She discusses that since biased input data is the most common reason given for this negative relationship, the problem must be tackled at its core. This core consists of the fact that humans are the ones who put in the input data and select the criteria on which an algorithm needs to find the best candidates. When these employees are more aware of biases that could occur within this process, but also biases in general, they will be more conscious of which criteria they will select and what data they put in the algorithms. According to the interviewee, with this HR-intervention you not only tackle this specific problem, but also enhance more diversity within the organization in the end. This point of view is a valuable addition to this literature review and for organizations in practice.

Appendix B

Summary Interview Interviewee 2

The second interview was held with an HR-advisor at DLA Piper, which is an international law firm. As an HR-advisor at DLA Piper the interviewee is responsible for the entire HR lifecycle of the junior advocates. This entails being the first point of contact for these employees, the inflow and outflow, performances, promotions etc. In addition, she also occasionally takes on some tasks within recruitment. DLA Piper does not make use of AI within the recruitment process. According to the interviewee, this is because law firms are not very progressive. Nevertheless, the interviewee does have a certain view on the research results I presented her, since she worked as a recruiter before she started working as an HR-advisor. Therefore, she knows how the recruitment process works, and to what you should pay attention to in this process. At DLA Piper much attention is paid to preventing any form of gender bias in the workplace. They have a diversity group for instance, which is a group of employees who focus on diversity regarding gender. This group organizes a women's week every quarter, in which they pay extra attention to the women in their company, since most of their workforce consists of men. In this week, they let different female employees of DLA Piper tell a success story of themselves to the rest of the company.

The interviewee's view on the research results regarding the relationship between AI recruitment and gender bias gave valuable insights. The interviewee stresses the need for an awareness training for older personnel as an intervention to prevent gender bias in AI recruitment. In her opinion, senior staff focuses more on the content of their tasks rather than being aware of possible ethical concerns that come along with it. The reason the interviewee gives for this, is that she thinks that students nowadays get more educated about topics like gender diversity then the older generations got. Therefore, she thinks that the younger generation, who recently graduated, is more aware of ethical concerns like gender bias.

Older staff that concerns themselves over the input of data into AI algorithms for recruitment processes for instance, will be more aware of gender biases after these kinds of trainings. With this intervention, the interviewee tries to tackle the most common reason the literature gives on why AI recruitment and gender bias are negatively related to each other, namely the biased input data. The interviewee also mentioned that this intervention could be implemented at DLA Piper when they might start working with AI in the recruitment process in the future, as they do not work with AI recruitment right now. Therefore, the insight the interviewee gave regarding this intervention could be a relevant practical implication for organizations who are already using AI in their recruitment processes to mitigate or even remove gender bias.