Artificial Intelligence in Hiring: A Critical Review of Bias in Recruitment Algorithms

F. P. Godoy Ojeda

SNR 2081928

Supervisor: A. Spanouli

Second assessor: K. Odstrčilíková

School of Social and Behavioral Sciences, Tilburg University

Human Resource studies, Bachelor’s thesis

May, 2025

Abstract

Artificial Intelligence (AI) has become increasingly prevalent in Human Resource Management, particularly in recruitment and selection processes. Organizations are adopting AI-powered tools to optimize hiring, improve efficiency, and enhance decision-making. However, the integration of AI in recruitment raises significant concerns regarding algorithmic bias, fairness, accountability, and transparency. This literature review analyzes fifteen peer-reviewed sources selected through predefined inclusion criteria to examine the relationship between AI-powered tools and bias. The review explores how bias can be introduced through unrepresentative or flawed training data, biased algorithmic design, and unexplainable decision-making processes. If not properly managed, AI systems may reinforce, reproduce, or even amplify existing human prejudices and systemic inequalities. The study identifies several mitigation strategies, including the development of fairness-aware algorithms, regular algorithmic audits, inclusive and representative data practices, and increased transparency in both model design and output. A recurring recommendation in the literature is the adoption of a hybrid recruitment model, wherein human judgment and contextual understanding work alongside AI technologies to ensure more equitable and sensitive outcomes. These findings highlight the complex ethical, legal, and operational challenges faced by recruiters and HR professionals when deploying AI tools. They also emphasize the importance of multidisciplinary collaboration among HR experts, data scientists, ethicists, and legal advisors. While AI offers notable advantages such as increased efficiency, scalability, and data-driven insights, its ethical implications demand rigorous oversight and continuous evaluation. Future research should further investigate the dynamics of human-AI collaboration in recruitment to support the development of more inclusive, transparent, and accountable hiring frameworks.

Keywords: Artificial Intelligence, Hiring bias, Algorithmic bias, Human Resource Management, Recruitment, AI-powered systems.

Artificial Intelligence in Hiring: A Critical Review of Bias in Recruitment Algorithms

Artificial Intelligence has been a topic of discussion since 1950. In recent years AI has evolved to a great extent and are being implemented in organisations as hiring tools, especially after the COVID-19 pandemic (O'Brien, 2024). The addition of these tools to the Human Resource departments has enabled recruiters to have an effective and automated recruitment process by reducing time-consuming tasks (Oswal et al., 2020). HR hiring tools are most commonly used in three steps: sourcing, screening, and interviewing or matching (O'Brien, 2024). During the process, the algorithm in hiring tools looks for similar traits or profiles of past applicants who are successful in the job, such as of the managers (Andrews & Bucher, 2022).

According to Langenkamp et al., (2020), there are four categories that can lead to algorithms failing: the dataset failure, the model structure, metrics, and the application of the model. It is known that traditional hiring methods bring bias along because people naturally make subjective decisions; however, since algorithms are developed by humans and trained on past data, it is often unclear whether they reduce bias or simply reproduce it in a less visible form (Harvis-Nazzario, 2022). This is what Langenkamp et al., (2020) calls the *dataset* failure. In short, a dataset failure happens when past data is used to train the algorithm as this data is often manipulated by humans and hence, bias is therefore not effectively eradicated in the hiring process. *Model structure* follows as machine learning has different methods for 'learning' and 'predicting'. More precisely, a model called a 'convolutional neural network' is excellent at recognising pictures, but this model will not work if the aim is to understand someone's voice; if the wrong model is used in AI hiring, this will result in mistaken decision making (2020). *Metrics* is the set of statistics behind the way the model performs by looking at values such as false positives, false negatives, and accuracy. In relation to AI hiring, metrics can include statistics related to fairness – aiming at making concepts such as ‘fairness’ and ‘equality’ measurable. At the same time, introducing fairness related numbers into the system ensures that the error rate for different demographics are the same (Langenkamp et al., 2020). Metrics cause algorithmic failure by turning complex societal issues into simple formulas as it ignores historical inequalities and hidden biases (Langenkamp et al., 2020). Finally, the *application* of the model justifies existing biases in the hiring system leading to unfair decision making (Langenkamp et al., 2020).

AI-powered recruitment tools and bias is a relationship that has practical and theoretical relevance. Although there is a lot of information circulating about AI hiring tools and bias, this literature review is an addition to the already existing research, aiming at focusing on bias in candidate selection instead of conducting a detailed analysis in a specific type of bias e.g., gender bias, age bias, race bias (Fabris et al., 2023). Furthermore, due to the rapid evolution of technology, many existing literature reviews have become dated, making it difficult to obtain current insights on the topic, available data might not be applicable nowadays as the models would have to be adapted accordingly (Mujtaba et al., 2024). As AI-powered hiring tools are now being widely and increasingly adopted by companies in their candidate selection process, it is essential that this literature review contributes to existing research by focusing on a more recent time frame—particularly from 2018 to the present— as the year 2018 is significant in the evolution of AI-driven hiring practices, thereby providing an updated perspective within the field (Fritts & Cabrera, 2021). The research contributes to the literature in algorithmic bias by exploring how AI systems inherit, amplify, or mitigate biases.

Practical relevance is also provided in this literature review as it is important for organisations to consider this for future AI-powered recruitment. According to Soleimani et al., (2022), knowledge sharing between HR managers and AI developers is an important factor to avoid bias in the hiring process. HR managers can help AI developers to avoid algorithmic bias to give a fair chance to candidates as the collaboration can lead to a better understanding of the AI-powered systems for recruiters. The use of this literature review provides an insight into improving fairness in hiring, by understanding if AI recruitment tools reduce or reinforce bias, organisations can opt for fairer hiring practices and hence, implement better decision making. Bias is often unintentional as the developers and users do not contemplate the discriminatory effects against some individuals (Coeckelbergh, 2020). Not only this literature review is an attempt to highlight the importance of team work between AI developers and HR professionals but also to raise awareness of the unintentional bias that is involved in selection processes with AI hiring.

As practical and theoretical relevance have been discussed, highlighting its importance, it is key to analyse the selected literature to further research this relationship. *"How do AI-powered recruitment tools relate to bias in candidate selection?"* is the question that is aimed to be answered with this research and will be answered after an in depth analysis of the literature.

Following, the theoretical framework will provide construct definitions and relevant background of "AI recruitment", "hiring bias" and for a better understanding "algorithmic bias" will also be defined. The scope of this literature review will be discussed as well in the following section. Later, the methods section, results, discussion and based on the analysis, a conclusion will sum up the research.

Theoretical framework

Considering AI is a growing field of study, many discussions have been made and various definitions have been used in past literature. Bias has been relevant for a long time and hence it has been addressed with a variety of definitions, including definitions for specific biases. This section will provide clear definitions for the constructs as theoretical perspectives that will help the analysis for the discussed relationship between AI recruitment and bias.

AI Recruitment

AI recruitment refers to the process of hiring through artificial intelligence where the AI tool makes choices based on data (Oman et al., 2024). AI examines huge quantities of information and distinguishes patterns and trends that might not be noticeable in traditional recruitment practices (Raghavan et al., 2020). According to Garg et al. (2021), AI hiring tools operate in three ways while conducting recruitment.

- *Sourcing:* it requires past data from professionals within the company that can be a match for the organisation looking at technical as well as soft skills. Involving the candidate's knowledge, academic background and experiences. If this is done correctly, there would be the right list of applicants for the job.
- *Screening:* this is the process of providing information and keywords to the system in order to match the applicant's resume to the keywords provided. In this step, applicants are filtered if their resumes do not contain the keywords that were provided to the AI system.
- *Matching*: this is done once the list is prepared for the suitable candidates. Matching consists of scanning the complete profile of the candidates that are suitable for the job with the required details such as salary, location and the core competencies. After, recruiters are left with a much smaller candidate pool and the interview stage begins.

AI-powered recruitment tools have proven to be useful for recruiters in order to find a sufficient number of the right candidates with the aimed qualifications and in an effective way and in consequence, finding the best-fit talent. It brings support when conducting difficult, repetitive and time-consuming tasks (Horodyski, 2023). However, AI recruitment presents some limitations that have to be considered as it has been a topic of discussion; bias.

Hiring Bias

In traditional hiring, it is common that HR employees are prone to cognitive biases unintentionally. This entails that there is a tendency for individuals to make systematic errors in their decisions and often leading to biases, for instance confirmation bias, halo effect, in-group bias and stereotyping bias (Derous et al., 2016; Kahneman & Tverskyv, 1979; Linos & Reindhard, 2015; Thomas & Reimann, 2023). Confirmation bias occurs when people search for information that supports what someone believes (Bratton, 2015, p. 344). The Halo effect is a type of cognitive bias where people's overall impression of a person, brand or things is influenced by one positive trait or characteristic, often leading them to assume other unrelated positive qualities (Bratton, 2015, p. 133). In-group bias, sometimes called in-group favouritism, is the tendency to favour the members of the team they belong to over the others (van Tubergen, 2020, p. 470). Although this bias is made by humans, it is important to understand what biases are to understand what algorithmic bias is.

Algorithmic Bias

Algorithmic bias is defined by Russell and Norvig (2021) as errors in AI systems that lead to inequitable outcomes for different groups. Algorithmic bias can also be seen as how societal relations influence the way technologies are shaped (MacKenzie & Wajcman, 1999). Algorithms learn bias through training data, amplified by algorithmic design or are able to arise from the ways systems are deployed (Kelan, 2023). As algorithms are created by humans, it is important to consider the unconscious cognitive bias that humans are prone to. Algorithmic bias is created depending on what data it is added to the system (Harvis-Nazzario, 2022). If the training data is biassed or not fully representative, the algorithm will reproduce inequitable results (Harvis-Nazzario, 2022). Overall, algorithmic bias is not only about flawed outputs but contextual, historical and societal dimensions plays a role in the way AI systems are built and used in hiring (Kelan, 2023).

It is essential to make a clear distinction between hiring bias and algorithmic bias. Hiring bias refers to stereotypes that are unconsciously created by human recruiters that influences their final decisions leading to unfair hiring practices and treatment for certain candidates. In contrast, algorithmic bias is caused when an AI system produces discriminatory outcomes as a result of flawed training data, biased design, or unintentional reinforcement of existing inequalities. Both types of bias lead to unequal opportunities but algorithmic bias can scale discrimination across thousands of applicants more rapidly and subtly. In summary, hiring bias is the unconscious outcome of an unfair decision made by a human recruiter, and algorithmic bias is an unfair decision based on the trained data that was used. The similarity is that both lead to discrimination of certain groups of applicants. The discrimination caused by the algorithm and by human recruitment includes but it is not limited to women discrimination, ethnic minorities discrimination, and discrimination towards people with disabilities (Fabris et al., 2023). Discrimination against women refers to unfair treatment of individuals based on their gender. Ethnic minorities face discrimination based on their ethnic background or cultural identity. Discrimination against individuals with disabilities results in prejudicial treatment of individuals with physical, mental, or sensory impairments, in doing so creating barriers to full participation in society (Fibbi et al., 2021).

Social Cognitive Theory

Badura (1986) has defined the social cognitive theory as the extent to which people control their behaviour and environment based on self-preservation. The Social Cognitive theory is a psychological framework explaining how people learn from their surroundings and from observing others, and how they use the information learned to make decisions and exhibit specific behaviours. This theory helps to understand people's cognitive and affective factors by considering their behavioural patterns and evaluating the environmental factors that influence their response. Factors including age, group size, gender, status, and socially assigned roles. These factors play an important role when a response is generated by an individual, hence bias is created (Bandura, 1999). Thoughts, feelings and actions are shaped not only by personal experience but also by societal context and roles the individuals take in society.

Not only age, group size, gender, status and socially assigned roles influence but also through mechanisms such as observational learning and social reinforcement, individuals internalize behaviours that align with dominant norms, including those that disadvantage certain groups. Observational learning means that people learn by watching actions and outcomes of others' behaviours and not only by experiences. Social reinforcement involves receiving approval or disapproval from others and this encourages or discourages particular behaviours (Bandura, 1999).

Companies use AI to screen job applicants with data that was previously entered in the system by humans based on their own beliefs, preferences and biases. As the data is entered by humans and as it was already mentioned, people's behaviours and decisions are shaped by social cognitive factors (environment, social roles, and observational learning), meaning that these unconscious biases get introduced into the AI system. An example of training data that includes self-preservation, refers to the human tendency to protect one's interests, status, or group members. Individuals involved in creating or using AI systems may unconsciously favour candidates who resemble themselves because from their perspective this is safer or more beneficial for maintaining their social position. The Social Cognitive Theory will support this literature review to understand the process behind individuals' reactions to the environment and hence, how the unconscious bias is entered into the system as data.

This literature review researches the relationship between artificial intelligence (AI) recruitment and bias. The research uses multidisciplinary sources with a focus on diversity and fairness by examining how AI is being a support tool for hiring practices, as well as the ethical implications and strategies in order to mitigate algorithmic bias.

Methods

Study Design

The research question was answered using a literature review on the relationship between AI recruitment tools and bias on candidate selection processes. A literature review is a secondary literature analysis that summarises existing research (Hempel, 2020). Hempel (2020) defines a literature review as standardised research methodology that synthesises existing evidence aiming at answering a research question by applying steps to reduce reviewers bias. This literature review will follow the Critical Appraised Topic (CAT) approach, meaning that this research will conduct a standardised summary of existing research evidence in order to address the research question and to evaluate the relevance of the findings (Sadigh et al., 2012). A detailed discussion of the methodology used to conduct this literature review follows, including the step-by-step debrief of how and why the literature was chosen, as well as the criteria that was applied to select the existing literature that will be analysed.

Search Strategy

In order to find relevant literature for this topic, the databases used were WorldCat and Google Scholar. The search was conducted between 26th of February up to and including 25th of March. As these two databases yielded a large volume of results, in order to filter available data that is not related to the topic of interest, keywords and boolean operators were entered in both databases (WorldCat and Google Scholar) and combined as follows: "AI hiring AND bias" as well as "AI AND recruitment bias" and "AI AND diversity and inclusion". Table 1 shows a clear overview of the literature presented once the combination of keywords and boolean operators were searched. As shown in Table 1: when using "AI hiring AND bias" the number of results shown in Google Scholar was n=154 and in WorldCat n=1,900. With the terms "AI AND recruitment bias" Google scholar reported n=605 results and in WorldCat n=3,900 results were shown. "AI hiring AND diversity and inclusion" showed a total of n=79,800 in Google Scholar and WorldCat showed n=456 results.

Table 1

Literature Search Terms and Boolean Operators Used

Search Terms	WorldCat	Google Scholar
"AI hiring AND bias"	1,900	154,000
"AI AND recruitment bias"	3,900	605,000
"AI hiring AND diversity and inclusion"	456	79,800

Selection Criteria

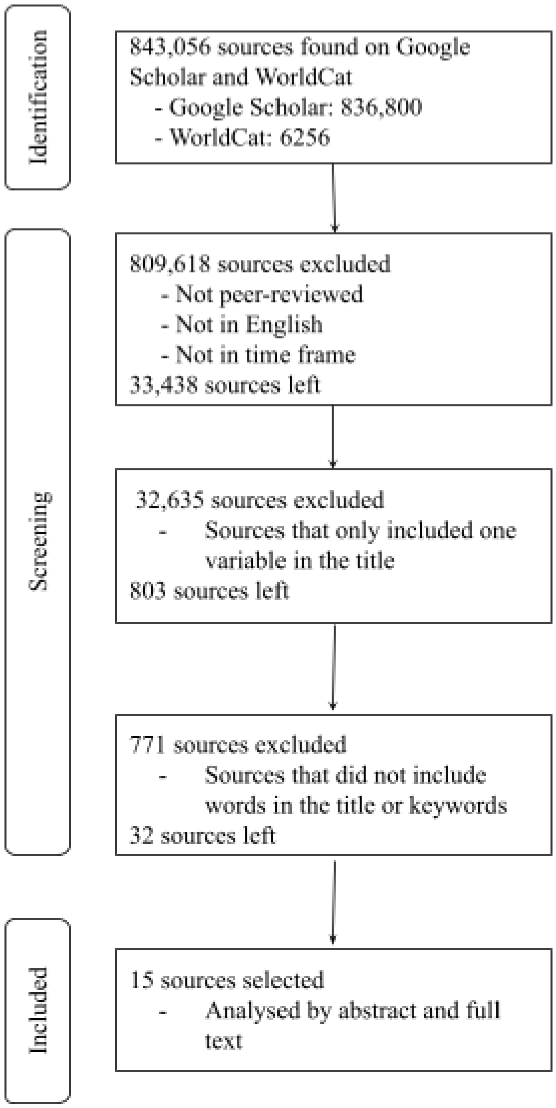
As Google Scholar and WorldCat presented a total of n=842,056 results, the need to filter articles that included unrelated topics was addressed by applying selection criteria. In order for existing literature to be considered, papers had to be written in English and peer reviewed. After this criterion was applied, there was n=745,572 excluded from the initial search between all the databases. As there was no country or region chosen for this review, there were no articles excluded by the category of country or region. Moreover, as artificial intelligence has been embraced by HR professionals since 2018 (Upadhyay, 2018), the chosen literature ranges between 2018 and 2025. Once this filter was applied, n=33,438 articles were quickly addressed by title looking for relevant keywords.

Screening & Selection Process

Because of the considerable number of articles still needed to be filtered, it was important that the title of the articles included the topic of interest; AI recruitment and bias. An approximation of n=32,635 articles were filtered based on title of relevance. The next screening stage was done with n=803 articles; the articles that had the most relevant titles and keywords such as hiring bias, AI bias, hiring tools, equity, inclusion, discrimination, artificial intelligence, fairness, AI ethics, and mitigation strategies were selected (n=32). After this criteria was applied, n=771 articles were excluded. The articles that met the title and keywords criteria (n=32) were analysed by their abstract and full text reaching to the final 15 sources that were selected as they provided valuable information that will help to answer the research question of this literature review.

Figure 1

PRISMA



From Raw to Analysable Data

In order to analyse the data from the selected literature, the text of the fifteen articles was carefully read and analysed to have a clear understanding of what is the main focus of each article. This initial step involved a thorough reading of each literature, taking detailed notes on core themes, objectives and outcomes to ensure full comprehension of the articles. Simultaneously, a findings table containing the author, keywords, type of research and a general overview of the main findings was included for each article. This table was systematically constructed to ensure consistency and coherence in the representation of data for all sources. The findings table can be found in Appendix A; aiming at summarizing the findings of each article to have a clear overview of the information presented in the analysed literature that will lead to the answer of the research question. The findings table in Appendix A serves as a structured tool for data organisation providing an easier comparison and synthesis of the literature.

After the findings table was concluded, the analysis was conducted in a thematic way, as three categories that were relevant to answer the research question were found. The thematic analysis involved identifying recurrent patterns and clustering them into categories that are present in the literature. The categories that were found and that the analysis will focus on are: Findings on the Absence of Bias in AI Systems, Findings on the Persistence of Bias in AI and Methods of Mitigating Bias While Using AI in Recruitment and Selection Processes. These themes emerged organically from the content of the selected articles and were not predetermined. This thematic structure facilitates a focused and in-depth discussion aligned with the goal of answering the research question.

Findings

Patterns found in the literature lead to answer the research question regarding the relationship of AI-powered hiring tools and bias. As AI is constantly evolving, the use of AI for hiring purposes grows accordingly, leading to organisations and scientific literature lagging behind the current state of technology (Cohen, 2019). The adoption of AI-powered recruitment tools has enhanced effectiveness in recruitment processes but conflict has risen between humans and machines, as the idea that machines will take human's jobs is often discussed (FraiJ & László, 2021). Even though it has been proven that machines have reshaped jobs and replaced humans in jobs, the idea of human-technology cooperation looks more appealing nowadays (Vivek, 2023).

Findings on the Absence of Bias in AI Systems

Considering the way algorithms are trained as recruiting tools, the studies by Cohen (2019), FraiJ and László (2021), Roumbanis (2025) and Vivek (2023) suggest that AI recruitment can reduce bias in the recruitment and selection processes. FraiJ and László (2021) reported that AI technologies can automate repetitive tasks, reduce human biases, and improve efficiency in recruitment leading to unbiased and rapid decision-making. Vendors such as HireVue, Retorio and myInterview are used nowadays in order to support organisations aiming for a more diverse workforce. These tools are capable of conducting fully automated video interviews that are blind to race/ethnicity, race, and gender (Roumbanis, 2025).

FraiJ and László (2021) reported that AI is intelligently designed so that name, age, gender, race, and belief can unbiasedly pass through AI systems. Roumbanis (2025) presents the term "meta-algorithmic judgements", this refers to the process of decision making that human recruiters go through when working side-by-side with AI. The decision making of the human recruiter is regulated by the organisation but suggestions provided by AI-powered systems are considered as well. After the interviews done by Roumbanis (2025), results suggest that the future AI recruitment will be automated but always with the support of human discretion; hybrid intelligences. This goes hand in hand with Cohen's (2019) and Vivek's (2023) ideas that AI can reduce bias when it is implemented correctly and with human oversight in order to support diversity initiatives by identifying and mitigating unconscious bias. Vivek (2023) specifically mentions that human oversight brings understanding, emotional intelligence and ethical considerations to the recruitment and selection process and hence, regular audits conducted by internal teams or third-party organisations specialised in AI ethics is beneficial for this process.

Findings on the Persistence of Bias in AI

As AI-powered tools have revolutionised recruitment, there is also a negative side to implementing AI tools in HR recruitment and selection processes. As FraiJ and László (2021) highlights "...an algorithm is only as good as the data on which it has been trained" (p.110). The problem with artificial intelligence is that it can be unreliable and is often bad at spotting subconscious bias – as well as humans – making it difficult to build an AI that checks for blind spots when the developers themselves are not aware of them (Cohen, 2019). Albaroudi et al., (2024) explain a bias that AI hiring can manifest during the recruitment process. For instance, measurement bias is present when AI fails to capture the skills and other traits relevant to the job. This means for example, if an organisation keeps hiring white candidates compared to Black applicants, AI will associate good performance with being white (Albaroudi et al., 2024).

The studies by Varsha (2023), Albaroudi et al., (2024), Tilmes (2022), and FraiJ and László (2021) mentioned the Amazon discrimination against women case as an example of how AI tools fail to represent the entire population. Amazon's algorithm was trained based on a decade of submitted CV's; predominantly CVs submitted by male applicants leading to AI-powered systems to prioritise male-centric-language patterns and discrimination against female applicants was seen (Albaroudi, Mansouri & Alameer, 2024). In addition, Tilmes (2022) researched discriminatory outcomes in AI hiring processes for applicants with disability. The results showed that criteria such as 'fit' and 'employability' legible to AI is not a neutral process as it is influenced by social attitudes and norms. An example of what AI relates to 'employability' is smooth speech, and regular eye contact, making employability a matter of body posture and that could bring about bias against people with a certain physical discapacity (Tilmes, 2022). Due to employers having less data on past or current disabled employees, AI hiring will not recognise or evaluate future disabled applicants as likely to be a good fit (Tilmes, 2022). This is an example of how underrepresentation of certain groups in datasets can influence the detection pattern of AI for matching processes.

AI hiring tools present gender and racial bias due to the data used in the training process of AI-powered hiring tools (Varsha, 2023). Moreover, the lack of transparency in AI models and missing or inadequate regulatory frameworks creates algorithmic bias due to training data resulting in AI amplifying societal inequalities (Albaroudi et al., 2024). In order to achieve inclusivity in AI hiring, more than algorithmic fairness is needed, as AI overlooks deeper issues resulting in bias (Tilmes, 2023).

Methods of Mitigating Bias While Using AI in Recruitment and Selection Processes

The selected literature proposes different strategies aiming at mitigating algorithmic bias. Bhatt (2023), agrees that AI is useful in the first stages of hiring; sourcing and screening. It suggested that the use of AI in later stages of the recruitment process is sensitive towards data privacy, transparency and defensibility issues (2023). Additionally, Bhatt (2023) and Albassam (2023) recognise that it is not about machine versus human recruitment but a cooperative recruitment process. Both articles presented human oversight as an important strategy to mitigate algorithmic bias and discrimination. AI is expected to cooperate with human capabilities of decision-making and selection of candidates instead of replacing human interventions (Bhatt, 2023).

Importantly, Ore and Sposato (2022) research concluded limitations on the use of AI not only for candidates but also for recruiters. AI-powered recruitment creates fear and distrust among recruiters as limitations regarding accuracy and reliability are in question (2022). In contrast, the articles by Bhatt (2023) and Albassam (2023) report that AI is a tool that provides recruiters with support not only for resume screening, predictive analytics, and virtual reality assessments but also by extending human capabilities in the selection of candidates. These authors suggest that it is of extreme importance that AI undergoes a proper human oversight process in order to mitigate algorithmic biases. If mitigation of bias is the aim, AI can support recruiters by making the process more effective and less time consuming and at the same time, recruitrers can support AI by evaluating flaws in order to make the process more fair for all candidates.

Bhatt (2023), Chan (2022) and Ferrara (2024) express the need for a deeper analysis in AI tools leading to a better understanding of how AI systems work and other organisational interests in order to be able to mitigate bias. Chan (2022) introduces the idea that fairness in AI should be a matter of procedural justice with close attention on how and why decisions are made. The findings are explained by the Equal Opportunity Merit Principle (EOMP) which suggests equal access to job opportunities for all eligible candidates taking into account ethical perspectives and hence, considering socio-economic realities (Chan, 2022). In addition to the EOMP, Chan (2022) suggests that Explainable AI (XAI) will enhance transparency, accountability and trust in AI tools from both parties, recruiters and candidates. XAI is described in Chan's (2022) article as an "human-interpretable description" for users regarding the factors leading to an AI-generated decision. Moreover, it is not only about comparing humans to AI but also about a close examination of AI systems by paying attention to strengths and weaknesses (Gao & Cheung, 2024; Albassam, 2023). Ferrara (2024) presents a mitigation strategy for bias data entered in AI systems through data collection, algorithm design and deployment. Solutions such as fairness-aware machine learning and transparent model designs to ensure representation of the entire population were presented (2024). As Chan (2022) and Ferrara (2024) mentions the importance of a deeper analysis of how and why AI decisions are made, Bhatt (2023) suggests that mitigating bias requires evaluation of other organisational interests on the use of AI-powered systems based on information security and return on investment (ROI) focusing on fairness on later stages of the recruitment process.

Findings suggest that mitigation of bias in AI recruitment involves not only HR employees but also other stakeholders (Gao & Cheung, 2024; Rigotti & Fischer-Villaronga 2024; Hunkenschroer & Luetge 2022). Organisational and policy level interventions are key in addressing bias making it managers responsibility to ensure a fair recruitment process. There is a need for clear regulations, anti-discrimination protections and ethical standards that have to be implemented to be able to have a bias-free recruitment process when using AI. The ethical standards will reduce the risk of unfairness, bias, accountability and transparency. The need for a better theoretical framework and empirical studies in order to understand how AI can enhance or reduce diversity, equity and inclusion (DEI) in the hiring processes is an important factor when talking about reducing algorithmic discrimination (2024).

Discussion

This literature review examined the relationship of AI-powered tools and bias in the recruitment process. By using data that was collected from fifteen sources chosen after applying a selection criteria, the aim of this research is to answer the following research question: *"How do AI-powered recruitment tools relate to bias in candidate selection?"*. The studies were analysed and findings were analysed in order to answer the research question.

Findings suggesting that AI is free of bias provided examples of AI tools that are used in organisations aiming at a more diverse workforce such as HireVue, Retorio and myInterview. These are examples of AI-powered systems that are capable of conducting video interviews that are blind to ethnicity, race and gender (Roumbanis, 2025). This means that there are AI systems that have proven to be non-discriminatory against ethnicity, race and gender. An explanation of why these AI systems are free of data can be provided by the results of Puyol-Antón's et al. (2021) study. This study introduces the concept of balanced data, defined as the process of adjusting the representation of different demographics or categorical groups in a dataset in order to ensure equitable training conditions in machine learning (2021). The study reported that AI systems containing balanced data can show unbiased results (2021). As AI bias has been found to be unbiased, this can be attributed to the data received by the algorithm.

Conversely, findings that showed that AI is related to bias was presented. Bias in AI systems is introduced by humans in the AI-system training process (Kelean, 2023). Measurement bias and representation bias are discussed by Albaroudi et al., 2024; highlighting how AI tools fail to capture the skills and traits that are important for the job; as well as a failure of representation of the entire population. The Amazon discrimination against women in the use of AI-systems during recruitment process was discussed in multiple articles as well as discrimination against disabled applicants (Albaroudi et al., 2024; LraiJ & László, 2021; Tilmes, 2022; Varsha, 2023). These examples have demonstrated that AI is related to bias in the recruitment process based on the training data received by the algorithm. The Social Cognitive Theory by Bandura (1999) explains the role of social context and learned patterns in shaping individual cognition and behaviour. The theory supports the explanation of why AI systems replicate human biases. It has been found that training data entered in the AI-powered systems are introduced by humans, meaning that the training data contains prejudices, stereotypes and systematic inequalities (Bandura, 1999 & Kelean, 2023). Using the perspective of Social Cognitive Theory, AI bias emerges as the AI systems imitate human social behaviour introduced in the data used to train the algorithm.

Furthermore, the last pattern that was recognised after an in depth analysis of the selected sources, is ways of mitigating bias during recruitment processes that involve AI-powered tools. Results showed that since bias is introduced by humans in the training process of AI systems, human oversight is key for a bias free recruitment process as humans bring factors that AI tools do not possess; the capability of understanding, emotional intelligence and ethical considerations (Vivek, 2023). Recruiters and third parties can be involved in the process of human oversight as a way to mitigate bias (Vivek, 2023). In addition, implementing policies and regulations that are anti-discrimination and regulating ethical standards are key to achieve a bias-free recruitment process (Rigotti & Fischer-Villaronga, 2024). As AI recruitment not only affects applicants but also recruiters; it is suggested that in order to mitigate bias it is important to consider the Equal Opportunity Merit Principle (EOMP) that considers socio-economic factors as well as ethical perspectives aiming at equality on the accessibility of job opportunities for all eligible candidates. The use of Explainable AI (XAI) will build trust in recruiters and applicants as it encourages transparency, accountability and trust in AI tools (Chan, 2022). Finally, in order to mitigate bias Gao & Cheung (2024) suggests that it is necessary to have a better theoretical framework and empirical studies to understand the role that AI plays when promoting or hindering diversity, equity and inclusion as there are currently inconsistencies in existing literature.

Artificial intelligence is related to bias even though there is proof that existing vendors use bias free systems for recruitment as these can contain balanced data. As bias is introduced by humans, the mitigation strategies suggested in this research are important to consider as well as the cooperation between human recruitment and AI recruitment. The idea of hybrid intelligence should be appealing to recruiters as it is important the cooperation between both in order to get bias-free results on the recruitment and selection process. AI is a persistent challenge as AI-powered systems relate to bias and factors such as proper training, mitigation strategies and more research will only reduce bias.

Limitations

This study is a literature review aiming at synthesising existing research regarding the relationship of AI-powered systems and bias. This literature review presented the definition of the terms ‘hiring bias’ and ‘algorithmic bias’, however, the analysed articles contain different definitions. As there is no universally agreed definition of ‘bias’ in AI, the definition of ‘bias’ in AI can include data bias, algorithmic bias, and societal bias. When studies use inconsistent definitions of key concepts, the challenge to make meaningful comparisons within the studies is present as different definitions often lead to different methodologies and this weakens the reliability of the study.

Concurrently, this review focuses on the analysis of past reviews without the inclusion of primary data. This means that this study fails to validate the findings that are reported as it relies on data quality and interpretations of other researchers. The dependency on data quality and interpretations of other studies is a limitation as these can contain error or bias. At the same time, this literature review does not generate new knowledge limiting the value in exploring unexplored dimensions.

This literature review selected sources using strict criteria, this means that grey literature (literature published outside of academic or commercial channels) reports were filtered out. This is a limitation as it narrows the diversity of perspectives regarding ethical standpoints. Language is a limitation for this study, as only English articles were analysed leading to this literature review to restrict diversity. This is a limitation as there is important information in that this literature is missing as well as methodologies from non-English-speaking articles and this could mislead findings into skewed conclusions, under representation of regional studies and limited applicability.

Future research recommendations

After the sources have been analysed, it is clear there is a need for a deeper understanding of AI systems bias on marginalised communities. The importance of understanding the long-term consequences of AI-powered systems and bias in these communities will lead to a better understanding of inequality. Topics on how social mobility, opportunities, and access to resources is related to bias is an open field of study (Eubanks, 2017). Subsequently, many studies focus on theoretical intervention strategies for bias mitigation but there is space for future research to test these strategies and evaluate their effectiveness once they have been applied in organisational practices. In order to identify the best context-specific best practices, it is important to evaluate the mitigation strategies for AI systems and bias not only in hiring but in different fields such as healthcare and criminal justice. Finally, future research should stop presenting findings as machine versus humans but instead focus on the cooperation of AI and human recruitment (Albaroudi et al., 2024). To illustrate, AI can be useful for *sourcing* and *screening* by helping the recruiter to make these processes more efficient but for the *matching* stage, human recruiters and AI can work together in order to balance the algorithm with emotional intelligence and understanding. This allows future research to investigate how successful the term ‘hybrid intelligence’ can be as a new recruitment and selection process strategy.

Theoretical and practical implications

After analysing the findings, this research contributes to theory in different ways. Firstly, this study supports theories regarding bias in AI-powered systems by defining algorithmic bias and giving a better understanding of how algorithms are trained and hence, how bias is reflected in the candidate selection process involving AI. At the same time, AI is a growing field and it is difficult to keep up with the rapid growth of technology. For this reason, this study adds to theory by contributing with a more recent study providing a newer form of information. Finally, this study supports Roumbanis' (2025) idea of future hiring processes where hybrid intelligence is used as a way to mitigate bias, where AI and human recruitment co-work to achieve fairness and bias-free recruitment processes by suggesting this is the future of hiring.

Moreover, the practical contribution of this study is that by understanding how AI causes bias in the hiring process, organisations can use the mitigation strategies from the findings of this paper. An example of this is the above mention of hybrid intelligence (Roumbanis, 2025) as well as the Explainable AI (XAI) method and the Equal Opportunity Merit Principle (EOMP) guiding organisations while the use of AI-powered systems (Chan, 2022). As mentioned in Appendix B, interviewees mentioned the importance of human oversight and agreed that hybrid intelligence should be used in hiring in order to reduce the risk of discrimination. At the same time, if companies use hybrid intelligence, they are less exposed to hire someone who does not fit into the culture of the company or the team (Appdendix B, interviewee A). Furthermore, this study gives positive and negative insights of AI hiring practices but the negative use can grow concerns about algorithmic discrimination and hence, guide policymakers and HR professionals in creating stricter guidelines for a responsible use of AI during recruitment processes.

Conclusion

Ultimately, this research highlights the positive and negative relationship of AI-powered systems and bias but also ways in which the negative relationship can be mitigated. The study provides understanding on how bias is replicated by the algorithm as well as AI systems that are currently used and have not shown bias. Findings conclude that there are multiple strategies to mitigate algorithmic bias suggesting that it is not about humans versus machines but a cooperation between both sides providing their knowledge and skills aiming at the same outcome; bias-free recruitment process.

Reference List

Albaroudi, E., Mansouri, T., & Alameer, A. (2024). A Comprehensive Review of AI Techniques for Addressing Algorithmic Bias in Job Hiring. *AI* , *5* (1), 383–404. https://doi.org/10.3390/ai5010019

Andrews, L., & Bucher, H. (2022). *Automating Discrimination: AI Hiring Practices and Gender Inequality | Cardozo Law Review*. https://cardozolawreview.com/automating-discrimination-ai-hiring-practices-and-gender-inequality/

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), e02089. https://doi.org/10.26668/businessreview/2023.v8i6.2089

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

Barends, E., Rousseau, D. M., & Briner, R. B. (Eds.). (2017). *CEBM a guideline for critically appraised topics in management and organizations* (Version 1.1). Center for Evidence-Based Management.

Bhatt, P. (2023). AI adoption in the hiring process – important criteria and extent of AI adoption. *Foresight, 25*(1), 144–163. https://doi.org/10.1108/FS-07-2021-0144

Bratton, J. (2015). *Introduction to Work and Organizational Behaviour* (Third Edition). Palgrave. DOI 10.1007/978-1-137-43206-3

Cohen, T. (2019). How to leverage artificial intelligence to meet your diversity goals. *Strategic HR Review, 18*(2), 62–65. https://doi.org/10.1108/SHR-12-2018-0105

Coeckelbergh, M. (2020). *Bias and the meaning of life. In AI ethics* (Chapter 9). The MIT Press. https://doi.org/10.7551/mitpress/12549.003.0011

Chan, G. K. Y. (2022). AI employment decision-making: integrating the equal opportunity merit principle and explainable AI. *AI & SOCIETY: Journal of Knowledge, Culture and Communication, 39*(3), 1027–1038. https://doi.org/10.1007/s00146-022-01532-w

Derous, E., Buijsrogge, A., Roulin, N., & Duyck, W. (2016). Why your stigma isn't hired: A dual-process framework of interview bias. *Human Resource Management Review, 26*(2), 90–111. https://doi.org/10.1016/j.hrmr.2015.09.004

Eubanks, V. (2017). *Automating inequality: how high-tech tools profile, police, and punish the poor* (First edition). St. Martin's Press.

Fabris, E., Van Herck, P., Pušnik, M., Sokka, L., & Van Otterdijk, S. (2023). *Fairness and bias in algorithmic hiring: A multidisciplinary survey*. arXiv. https://arxiv.org/pdf/2309.13933

Ferrara, E. (2024). *Fairness and Bias in Artificial Intelligence: A Brief Survey of Sources, Impacts, and Mitigation Strategies. Sci, 6*(1), 3. https://doi.org/10.3390/sci6010003

Fibbi, R., Midtbøen, A. H., & Simon, P. (2021). Concepts of discrimination. In R. Fibbi, A. H. Midtbøen, & P. Simon (Eds.), *Migration and discrimination* (pp. 13–20). Springer. https://doi.org/10.1007/978-3-030-67281-2_2

FraiJ, J., & László, V. (2021). A Literature Review: Artificial Intelligence Impact on the Recruitment Process. *International Journal of Engineering and Management Sciences, 6*(1), 108-119. https://doi.org/10.21791/IJEMS.2021.1.10

Fritts, M., & Cabrera, F. (2021). AI recruitment algorithms and the dehumanization problem. *Ethics and Information Technology*, *23*(1), 1–11. https://doi.org/10.1007/s10676-021-09615-w

Gao, G. Y., Lyu, J., & Cheung, C. M. (2024). *AI Recruiting and Workplace Diversity, Equity, and Inclusion: A Literature Analysis*.

Garg, A., Gaur, S., & Sharma, P. (2021). A review paper: Role of artificial intelligence in recruitment process. *Anwesh: International Journal of Management & Information Technology, 6*(1), 33–37.

Geetha, R., & Bhanu, S. R. D. (2018). Recruitment through artificial intelligence: a conceptual study. *International Journal of Mechanical Engineering and Technology, 9*(7), 63-70.

Geller, E. S. (2016). *The Psychology of Safety Handbook*. (n.p.): CRC Press.

Harvis-Nazzario, L. (2022). It's not the algorithms, it's the people: Preventing bias in automated hiring tools starts with humans. *Rutgers Computer and Technology Law Journal, 49*(1), 138–175.

Horodyski, P. (2023). Recruiter's perception of artificial intelligence (AI)-based tools in recruitment. *Computers in Human Behavior Reports, 10*, Article 100298. https://doi.org/10.1016/j.chbr.2023.100298

Hunkenschroer, A. L., & Luetge, C. (2022). Ethics of AI-Enabled Recruiting and Selection: A Review and Research Agenda. *Journal of Business Ethics, 178*(4), 977–1007. https://doi.org/10.1007/s10551-022-05049-6

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*(2), 263–291. https://doi.org/10.2307/1914185

Kelan, E. K. (2023). Algorithmic inclusion: Shaping the predictive algorithms of artificial intelligence in hiring. *Human Resource Management Journal, 34*(3), 694–707. https://doi.org/10.1111/1748-8583.12511

Langenkamp, M., Costa, A., & Cheung, C. (2020). *Hiring fairly in the age of algorithms*. *arXiv*. https://doi.org/10.48550/arXiv.2004.07132

Linos, E., & Reinhard, J. (2015). *A head for hiring: The behavioral science of recruitment and selection*. Chartered Institute of Personnel and Development.

MacKenzie, D., & Wajcman, J. (1999). *The social shaping of technology*. Open University Press.

Mujtaba, D. F., & Mahapatra, N. R. (2024). *Fairness in AI-driven recruitment: Challenges, metrics, methods, and future directions*. *arXiv*. https://doi.org/10.48550/arXiv.2405.19699

O'Brien, T. (2024). When machines make hiring decisions: Examining the risks and limitations of AI-based recruitment tools. *Florida State University Law Review Online, 51*, 20-37. https://heinonline.org/HOL/Page?collection=journals&handle=hein.journals/fsuon51&id=23

Oman, N. Z. U., Siddiqua, N. A., & Noorain, N. R. (2024). Artificial Intelligence and its ability to reduce recruitment bias. *World Journal of Advanced Research and Reviews, 24*(1), 551-564.

Ore, O., & Sposato, M. (2022). Opportunities and risks of artificial intelligence in recruitment and selection. *International Journal of Organizational Analysis, 30*(6), 1771–1782. https://doi.org/10.1108/IJOA-07-2020-2291

Oswal, N., Khaleeli, M., & Alarmoti, A. (2020). Recruitment in the era of industry 4.0: Use of artificial intelligence in recruitment and its impact. *PalArch's Journal of Archaeology of Egypt/Egyptology, 17*(8), 39–47.

Puyol-Antón, E., Ruijsink, B., Piechnik, S. K., Neubauer, S., Petersen, S. E., Razavi, R., & King, A. P. (2021). *Fairness in cardiac MR image analysis: An investigation of bias due to* *data imbalance in deep learning based segmentation*. arXiv. https://doi.org/10.48550/arXiv.2106.12387

Raghavan, M., Barocas, S., Kleinberg, J., & Levy, K. (2020). Mitigating bias in algorithmic hiring: Evaluating claims and practices. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (FAT)* (pp. 469–481). Association for Computing Machinery. https://doi.org/10.2139/ssrn.3408010

Rigotti, C., & Fosch-Villaronga, E. (2024). Fairness, AI & recruitment. *Computer Law & Security Review: The International Journal of Technology Law and Practice, 53*. https://doi.org/10.1016/j.clsr.2024.105966

Russell, S., & Norvig, P. (2021). *Artificial intelligence: A modern approach* (4th ed.). Pearson.

Soleimani, M., Intezari, A., & Pauleen, D. J. (2022). Mitigating Cognitive Biases in Developing AI-Assisted Recruitment Systems: A Knowledge-Sharing Approach. *International Journal of Knowledge Management (IJKM), 18*(1), 1-18. https://doi.org/10.4018/IJKM.290022

Thomas, O., & Reimann, O. (2023). The bias blind spot among HR employees in hiring decisions. *German Journal of Human Resource Management, 37*(1), 5–22. https://doi.org/10.1177/23970022221094523

Tilmes, N. (2022). Disability, fairness, and algorithmic bias in AI recruitment. *Ethics and Information Technology, 24*(2). https://doi.org/10.1007/s10676-022-09633-2

Upadhyay, A.K. and Khandelwal, K. (2018), "Applying artificial intelligence: implications for recruitment", *Strategic HR Review, Vol. 17 No. 5*, pp. 255-258. https://doi.org/10.1108/SHR-07-2018-0051

Varsha, P. S. (2023). How can we manage biases in artificial intelligence systems – A systematic literature review. *International Journal of Information Management Data Insights, 3*(1), 100165. https://doi.org/10.1016/j.jjimei.2023.100165

Vivek, R. (2023). Enhancing diversity and reducing bias in recruitment through AI: A review of strategies and challenges. *Информатика. Экономика. Управление / Informatics. Economics. Management, 2*(4), 0101-0118. https://doi.org/10.47813/2782-5280-2023-2-4-0101-0118

Appendix

Appendix A: Findings table

Author (s) and Year	Keywords	Type of research	Main Findings
Albaroudi et al., 2024	Algorithmic bias, Deep learning, Curriculum vitae screening, Natural language processing, Artificial intelligence	Qualitative	Algorithmic bias is often due to training data, lack of transparency in AI models, and inadequate regulatory frameworks. In order to mitigate bias, some strategies include fairness-aware machine learning, bias mitigation algorithms, and transparent model design. Interdisciplinary collaboration is needed when developing ethical AI hiring processes.
Albassam, 2023	HRM technology, Artificial intelligence, Recruitment, AI-based recruitment strategies, Resume screening, Candidate matching, Video Interviewing, Chatbots, Predictive analytics, Gamification, Virtual reality assessments, Social media screening, Ethics, Legal standards	Qualitative	Resume screening, video interviewing, predictive analytics and chatbots, gamification, candidate matching, virtual reality assessments and social media screening tools are of great use for organisations including improved efficiency, cost savings, and better quality hires. On the other hand AI hiring raises concerns regarding algorithmic bias and discrimination.
Bhatt, 2023	HR technology, Artificial Intelligence and hiring, AI adoption, Future of hiring, Technology and hiring, Talent acquisition	Empirical	Recruiters prioritise information security and return on investment when evaluating the utilisation of AI hiring tools. AI implementation is suitable at initial stages of hiring: the sourcing and screening stages. AI hiring has evolved since COVID-19 pandemic making AI hiring a common strategy for recruitment processes.

Chan, 2022	Artificial intelligence, Employment decision-making, Bias, Fairness, Equal opportunity, Merit, Explainable AI	Theoretical and Conceptual	Based on the *equal opportunity merit principle* (EOMP) and *explainable AI* (XAI), fairness should be about procedural justice — how and why behind decision making and not only outcomes. The combination of EOMP and XAI will lead to transparency and accountability in algorithmic decisions fostering candidates' trust in AI tools. AI hiring should be fair and explainable in order for organisations to balance efficiency with justice.
Cohen, 2019	Human Resource Management, Diversity, Talent management, Recruitment, Talent	Qualitative	AI can reduce hiring bias when correctly and thoughtfully implemented and by hand with human oversight in order to achieve ethical outcomes. AI provides support for diversity initiatives by identifying and mitigating unconscious bias.
Ferrara, 2024	Artificial intelligence, Bias, Fairness, Discrimination, Mitigation strategies	Qualitative	Bias can be entered into the AI systems through data collection, algorithm design, and deployment contexts leading to inequalities; mainly towards marginalised groups. Bias mitigation strategies are discussed as well as organisational and policy-level interventions aiming at promoting transparency and accountability in AI development and deployment.
FraiJ & László, 2021	Artificial intelligence, Recruitment process, Staffing, Sourcing of Candidates, Human bias, Candidate communication	Qualitative	AI technologies can automate repetitive tasks, reduce human biases, and improve efficiency in recruitment. AI is able to process data similar to human cognition and that leads to unbiased and rapid decision-making enabling professionals to focus on strategic functions within talent acquisition.

Gao & Cheung, 2024	AI recruiting, Diversity, Equity, Inclusion, DEI, Workplace, Literature analysis	Qualitative	As a result of dividing the research in (1) applicants, (2) decision-makers, (3) mixed perspectives and (4) algorithms; Gao and Cheung highlight the inconsistencies and fragmentation in existing literature with the urge to have a better theoretical framework and empirical studies to understand AI role when promoting or hindering DEI in hiring procedures.
Hunkenschroer & Luetge, 2022	Artificial intelligence, Algorithmic hiring, Employee selection, Ethical recruitment, Ethics of AI, Bias of AI	Qualitative	Identifies ethical opportunities and risks in AI recruitment such as fairness, bias, accountability and transparency. Highlighting the lack of empirical studies and the need for specific ethical guidelines and more comprehensive theoretical foundations.
Ore & Sposato, 2022	Technology, Technology management, Technological change, AI	Empirical	The implementation of AI technology in recruitment and selection processes is also fraught with risks causing fear and distrust among recruiters.
Rigotti & Fischer-Villaronga, 2024	Fairness, Hiring, Recruitment, Discrimination, Data protection, Artificial intelligence	Qualitative	Applicants and employers have a different focus regarding fairness; employers focus on finding the best fit whereas HR practitioners highlight the importance of procedural fairness. Rules and regulations on the use of AI have to be addressed on anti-discrimination and data protection.
Roumbanis, 2025	Artificial intelligence (AI), Human judgement and decision-making, Hiring and personnel selection, Expert recruiters, Expectations and imagined futures, Sociology of algorithms	Qualitative	AI can help to create unbiased processes but with the cooperation of human-AI symbiosis where human judgment is key. The future of AI is imagined as increasingly automated but always with human discretion playing a crucial role within the process.

Tilmes, 2022	Artificial Intelligence, Fairness, Disability, Bias, Justice.	Qualitative and theoretical	In order to achieve inclusivity in AI recruitment, there is the need for more than only algorithmic fairness. A deeper understanding of complex realities such as disability injustices has to be recognised.
Varsha, 2023	Artificial Intelligence, Bias, Vulnerabilities, Responsible AI, AI ethics, AI systems.	Qualitative	AI tools in recruitment tend to express and amplify societal inequalities, AI hiring bias has demonstrated gender, and racial bias, favouring male and white applicants due to training data.
Vivek, 2023	Recruitment landscape, Technological advancements, modern recruitment, AI-driven recruitment, Unbiased decision-making	Qualitative	AI hiring could mitigate bias and enhance diversity but it might also replicate existing biases if it is not designed and monitored in a proper way. In order to do this, transparency, regular audits and human oversight is required as well as an interdisciplinary collaboration to ensure fair and inclusive outcomes when hiring.

Appendix B: Interviews

The findings of this study were discussed with two interviewees in order to have insights on practical implications. Both interviewees belong to the same department but in order to maintain their anonymity they will be addressed as interviewee A and interviewee B. Both interviewees work at Applied Medical at the tender and offers department. This department consists of 42 people and it is divided into different countries. The interview was conducted to the manager of the Spanish team and a team member of the same team.

Interviewee A

The interviewee demonstrated a general awareness of AI hiring, and acknowledged the growing precedes in recruitment and selection processes. She views AI as a powerful tool that benefits both recruiters and candidates, particularly making administrative tasks more efficient. However, she highlighted that AI should not replace human recruiters in hiring. It was mentioned that in-person interviews and personalised engagement remains essential for understanding the unique qualities of each candidate. She also highlighted the importance of how human interaction allows interveners to assess emotional intelligence, cultural fit, and interpersonal attributes; these are factors that algorithms may overlook. She argued that as AI can support decision-making, it cannot capture the full spectrum of human behaviour and motivation that is critical in team integration.

When asked regarding bias in algorithms, the interviewee was sceptical about the statement that AI is neutral. She recognised that as AI is being trained by humans and the data is replicated and amplified by the algorithm if it is not carefully managed. She mentioned that AI is capable and has the potential of reducing bias by applying consistent criteria and its effectiveness depends entirely on the quality and diversity entered during the training process. She expressed that in her opinion, in order for AI to be fair, AI should not only focus on technical qualifications but also incorporate personal and cultural dimensions when assessing applicants. Even though the company is not involved with AI in hiring, she has expressed concern for greater attention to algorithm design and the values embedded within it.

As the findings of this study introduced the term *hybrid intelligence,* the interviewee was asked regarding the opinion in this finding. She fully supports the use of hybrid intelligence where AI addresses technical qualifications and skills and human recruiters oversee the emotional and ethical dimensions. She believes collaborative approach is especially beneficial for roles that require strong interpersonal skills or cultural alignment like her team does. She believes that human oversight and policy reinforcement are seen as potentially effective strategies for bias mitigation or reduction, nonetheless she would remain cautious about relying on these solutions entirely. She expressed that AI can be used in both ways, and candidates can apply using fabricated CVs generated by AI tools, hence there is a need for regulation. She expressed that based on her experiences on her team, hiring an applicant based on AI can negatively affect team dynamics, prolong training, and reduce performance if candidates are selected based on flawed or narrow criteria. She views AI tools as a supportive tool rather than a replacement, manifesting the need for a balanced approach that preserves the human element.

Interviewee B

Interviewee B acknowledged having heard of the term “AI hiring” but admits to not being familiar with the processes or the algorithms involved. She believes AI is only as neutral as the data that it is trained on, emphasising that much of the data is created by humans whose perspectives are shaped by social, cultural, and historical influences. She believes AI systems may unintentionally reflect or even amplify these biases. In contrast, she believes that neutrality can be reached if AI is fed diverse and inclusive datasets spanning different contexts and time periods, arguing that systems are still too limited in scope to make this claim. She thinks AI has the potential to reduce bias in hiring but only if systems are carefully designed and implemented. At the same time, she manifests that human interaction in hiring is irreplaceable as personal conversations allow recruiters to assess values, character, and fit in ways that AI alone cannot.

As well as interviewee A, interviewee B was asked about bias in AI hiring tools. The interviewee expresses that AI systems can perpetuate and amplify societal biases, especially because historical inequalities in hiring have been encoded into the data that is used to train algorithms in a very successful way. She added factors as discrimination based on gender, race, age, class and other factors and expressed that even though there are companies working to overcome these biases, there are still major challenges that are present. At the same time, she expressed concern regarding the potential for AI to unintentionally discriminate hence the importance of preventive strategies. Interviewee B, suggested approaches for bias mitigation that included diversifying training datasets, regularly testing systems for bias, enforcing transparency, and maintaining constant human oversight.

When asked regarding the mitigation strategies suggested in this study, the interview expressed support for human oversight and policy enforcement, noting that while potentially effective, these are still vulnerable to human bias. For her, it is important to maintain active awareness of and objectivity in order to prevent historical biases from influencing decisions. She also supported the idea of hybrid intelligence, and expressed that it is an ideal strategy for balancing efficiency and fairness in recruitment. To add, ethical standards were mentioned as well as regulatory frameworks expressing that AI should be trained on credible, scientifically validated sources rather than unchecked online content. She admitted being aware that bias can never be completely eliminated, she believes that it can be significantly reduced with thoughtful design, transparent practices and vigilant monitoring. The interviewee expressed the usefulness of AI tools as a support for recruitment and it is even better if it is ethically developed and carefully supervised, not to replace humans but to aim at fair and inclusive hiring.