LUNDS UNIVERSITET

Ekonomihögskolan

Institutionen för informatik

AI and the Future of IT Recruitment

Addressing Bias within Hiring Processes

Bachelor Thesis 15 hp, course SYSK16 in Information Systems

Authors: John Arvidsson
Ludvig Eslind

Supervisor: Blerim Emruli

Correcting Teachers: Markus Lahtinen
Niki Chatzipanagiotou

AI and the Future of IT Recruitment: Navigating Ethical Concerns

AUTHORS: John Arvidsson, Ludvig Eslind

PUBLISHER: Institutionen för informatik, Ekonomihögskolan, Lunds universitet

EXAMINATOR: Osama Mansour, PhD

SUBMITTED: August, 2023

DOCUMENT TYPE: Bachelor Thesis

PAGE AMOUNT: 61

KEY WORDS: Artificial Intelligence, Discrimination, Bias, Recruitment, IT,

SUMMARY(MAX. 200 WORDS):
The use of artificial intelligence tools is rapidly increasing. Companies are now looking to use AI for recruitment in order to try to ensure a more fair process. This thesis aims to answer the question "How do recruiters in the Swedish IT sector perceive the effectiveness of AI tools in mitigating bias within the recruitment process?". While some recruiters see AI as a tool to increase objectivity and fairness in applicant selection, others are cautious believing the lack of transparency and explainability makes AI an unreliable resource. Previous literature includes the prevalence of bias in recruitment, and explains the difference in bias between different AI tools, but does not elaborate much of its perceived value by the human users themselves, a large component in its practical application and value. To answer this question, a quantitative analysis in the form of a questionnaire was conducted, and the responses analyzed. Results show that recruiters trust in their own ability to identify bias, but not the ability of others, pointing to human factors being the main reason for the emergence of bias in recruiting and AI. AI tools are perceived as slightly more trustworthy, but are not deemed effective in their current state.

Table of Contents

Figures

Tables

1 Introduction

The use of Artificial Intelligence (AI) tools has increased dramatically, and with the release of public tools such as ChatGPT, this is likely to continue. These AI tools now play a role in the job market itself, including the recruitment industry. The tools aim to help with recruiting for the companies that use them, but several concerns may arise as a result. One major factor is the question of bias, how does bias arise in AI tools, and how does it affect recruitment in the IT industry?

Traces of using AI tools, or in this case ChatGPT is being recognised across a large portion of hiring processes in companies today already. A 2023 study by ResumeBuilder shows that 97% of respondents were already using ChatGPT as part of their recruitment process, to some extent. The primary ways AI is being used for recruitment includes Writing job descriptions, drafting interview questions and responding to applicants.

However, several historical cases have been made where AI tools used are unfairly balanced. Najibi (2020) explains that racial discrimination in face recognition algorithms was apparent in the 2018 “Gender Shades” project, and incidents where AI was biased specifically in recruitment software has also been found, such as the secret 2016 Amazon recruiting tool, which showed substantial bias against female applicants (Dastin, 2018).

Previous concerns with having artificial intelligence tools has been the risk of profiling and bias based on insufficient databases or algorithms, which in some cases has made using AI recruitment tools less of a benefit and more of a concern. One example of this, brought up in the study “Ethical AI in facial expression analysis: racial bias”,, is that some ”state-of-the-art facial expression recognition methods such as Deep Emotion…” (Sham et al., 2022, p. 399) are bound to fail, since there is a “possibility of the methods being biased toward certain races. Therefore, a concern about fairness arises, and the lack of research aimed at investigating racial bias only increases the concern.” (Sham et al., 2022, p. 399). In this research, it was discovered that facial recognition systems were based on databases with a lack of balance between races in the data. This imbalance was then further magnified when the systems were put to use. Furthermore, the research shows that when data is balanced equally across different races, unbiased performance of the AI is obtainable (Sham et al., 2022).

The new AI tools that are being developed around the market may reduce these issues, but for AI tools to be used in recruitment efficiently, an understanding of the attitude of the user needs to be attained, in order to ensure an efficient human-machine collaboration. Therefore, this thesis aims to investigate the bias concerns of AI tools, and the general attitude of recruiters towards them.

1.1 Research Question

How do recruiters in the Swedish IT sector perceive the effectiveness of AI tools in mitigating bias within the recruitment process?

1.2 Purpose

The purpose of this study is to better understand how recruiters perceive AI tools and their ability to mitigate bias as part of the recruitment process in Swedish IT companies. This aims to provide further understanding of how artificial intelligence tools can be utilized to improve selection procedures, while also keeping in mind the potential risks they present.

This could provide companies and recruiters with useful information for future decision making when maintaining their recruitment processes, taking their attitude towards AI tools and its usefulness into consideration.

1.3 Boundaries

To keep the data relevant, boundaries have been set to address IT Companies within Sweden, as well as companies who recruit for positions for the Swedish IT job market. Although there are other ethical concerns with the use of AI such as Data Handling, Bias has been selected as the issue of investigation, as it is both relevant and measurable, while keeping the study focus clear and the scope relevant.

2 Literature Review

This chapter, divided into six parts, reviews previous literature to provide a basic understanding of the subject. It commences by explaining the recruitment process, AI, and bias. Next, it delves into bias within current recruitment practices, followed by its emergence in AI systems. Finally the chapter weighs the balance of trust and fear in AI interactions, and ends in a Thematic overview.

2.1 The Recruitment Process

Recruitment is defined as "the process of finding people to work for a company or become a new member of an organization" (Cambridge Dictionary, N.D.). The process of recruiting is divided into seven different parts, listed below.

2.1.1 Identifying the needs of the company

Before applicants themselves are part of the process, a company must first decide their needs. These needs include the exact position that needs to be filled, writing the title of the position, how it fits into the organization and the main duties. Both required as well as potentially meriting competences are also outlined here (Cook, 2009, p.92):

- "Very specific skills or knowledge that workers will acquire as part of their training, but would not possess beforehand, e.g. knowing how to serve meals and drinks on an aircraft."
- "More generalized skills or knowledge that organization might wish to select for, e.g. communicating well in writing."
- "Aptitudes that would make it easier for a person to acquire more specific competences, e.g. flexibility, or ability to learn quickly"
- "Personality characteristics, e.g. resilience, tolerance"

These attributes make up the core of what a company needs to gather and process before making a job advertisement (Cook, 2009).

Outlining the job position that needs to be filled, and knowing exactly what to look for in applicants to the position, as well as communicating this efficiently through a job advertisement is what makes this stage the start of the candidate selection process within recruitment.

2.1.2 Sourcing candidates

Finding people suitable for the job starts with reaching out to potential applicants. This can be done via relocation within a company, but also through external recruiting. A company can source candidates in many ways, such as public or private agencies, job fairs or e-recruitment (Cook, 2009).

The goal of this step in the process is to make the job as known as possible in order to gather as many potential high quality candidates as possible.

2.1.3 Screening applicants

Applicant screening is where documents are reviewed. CV, Application forms and resumes of all applicants are analyzed. Screening can be done in other ways too, such as telephone screenings or through references to the applicant.

Information gathered at this stage ranges all the way from education level, work experience, personality traits and more. Pre-screening questions may also be sent out to applicants in order to get further information about applicants.

The information gathered from candidates is used to determine the most suitable candidates for the open position, so called “shortlisting”.

2.1.4 Shortlisting candidates

Shortlisting is the process of selecting the most suitable candidates for a given role. Companies may choose to call references to verify the information applicants provided in the previous step in order to get a better idea of the applicant (Cook, 2009). Their overall fit for the company is ranked using a 1-10 scale, with the best fits being part of the list, known as being “shortlisted”. This list is then used to know which candidates should be called for an interview, effectively sifting out applicants not fit for the role advertised.

Bias is a clear problem at this stage and links to ethical issues (Cook, 2009). Studies have found that for example, candidates with African American names are less likely to get shortlisted than applicants with caucasian names, at a rate of 0.67 to 1 respectively (Cook, 2009). Maximizing fairness in the application is therefore important, as the applicants ideally should only be judged by their ability to do the job and their fit into the company. Studies by Kethley and Terpstra (2005) also show that the shortlisting process is not fully fair, finding both gender and racially based complaints, along with other forms of potentially discriminating factors, such as age.

2.1.5 Interviews

Interviews are a core part of the selection process. It involves recruiters asking questions to applicants, with the purpose of determining their fit into the company and position advertised, as well as an opportunity for the company to see how applicants are as people (Cook, 2009). They may be conducted by one or more people from the company or recruitment firm. d

Cook explains that interviews "seek to assess different attributes of candidates.", the most common ones being "personality dimensions, especially conscientiousness, most frequently assessed (35%), followed by social skills (28%), and mental ability (16%)", with other less common ones being "... knowledge and skills (10%), interests and preferences (4%), organizational fit (3%) or physical attributes (4%)." (Cook, 2009, p 95).

2.1.6 Testing

Mental tests are a valuable tool for assessing candidates' cognitive abilities, and they can be used to measure various aspects such as verbal, numerical, and reasoning abilities (Cook, 2009). The use of these tests helps employers to identify the candidates who have the necessary skills to perform the job.

In addition, Cook emphasizes the importance of ensuring that the tests used are job-related and consistent with the job requirements. This means that employers should select tests that measure the skills and knowledge required for the job. For instance, for a data analyst position, the employer may use tests that measure the candidates' skills in data analysis, statistics, and programming.

Personality is a crucial aspect of the selection process. Cook (2009) notes that personality tests can provide valuable information about candidates' behaviors, values, and motivations. These tests can be used to assess whether candidates fit with the organization's culture and the job requirements.

The tests used need to be valid and reliable, so that they do not discriminate against any particular group of candidates. Additionally, reasonable accommodations should be provided to candidates with disabilities to ensure that they have equal access to the testing process (Cook, 2009).

After testing is completed, the results are typically used in conjunction with other information gathered during the recruitment process, such as resumes, interviews, and reference checks, to make informed hiring decisions. Testing provides current data that can help employers assess applicant's ability, assisting employers in identifying the most qualified candidates for the job opening. This increases the likelihood of making a successful hire.

2.1.7 Extending a job offer

Finally, if the applicant has been chosen and has passed the testing phase, a job offer is sent to the applicant by the employer or recruitment firm. If the applicant accepts this offer, a contract is to be signed, and applicants are given the starting date and other onboarding information, such as Terms and conditions, Salary, Benefits, etc. (Indeed, 2023).

2.2 Understanding AI, Machine Learning, and Algorithms

Artificial Intelligence (AI) is a field of computer science that aims to create machines that can perform tasks that typically require human intelligence, such as reasoning, learning, perception, and natural language processing. This field is interdisciplinary, drawing on concepts from computer science, mathematics, psychology, and philosophy (Russell & Norvig, 2010). The goal of AI is to create machines that can learn, reason, and act autonomously in complex and uncertain environments. One of the key features of AI is its ability to learn from data. This involves training algorithms on large datasets to identify patterns and make predictions or decisions without being explicitly programmed to do so.

Nilsson (1998) emphasizes that progress in AI requires a deep understanding of the nature of intelligence. He argues that intelligence is not a single, monolithic concept, but rather a collection of cognitive abilities that enable us to solve a wide range of problems. These cognitive abilities include perception, attention, memory, language, and reasoning. Developing AI is therefore not a simple question of progress in one field individually, but several combined.

Since AI is an umbrella term for a whole field, several subsets are highly relevant. However, in this section, focus will be on ML (Machine Learning) and DL (Deep Learning).

2.2.1 Machine Learning

Machine Learning is a subfield of artificial intelligence that involves the development of algorithms and models that enable computers to learn and make predictions or decisions based on data without being explicitly programmed (Murphy, 2012). The key idea behind ML is to use data to automatically discover patterns, relationships, and insights that can be used to make predictions or decisions on new, unseen data.

Jordan and Mitchell's (2015) article "Machine learning: Trends, perspectives, and prospects" provides a comprehensive overview of the field of machine learning. According to the authors, the goal of ML is to create algorithms and models that can learn from data, take action based on that data, and improve their performance over time. This is accomplished through a process of iteratively training the model on increasingly complex and varied datasets, with the aim of optimizing its accuracy and generalizability.

One key aspect of ML is its reliance on statistical techniques to extract patterns and insights from data. These techniques include regression analysis, clustering, and classification, among others. By applying these methods to large and complex datasets, ML algorithms can identify subtle relationships and correlations that would be difficult or impossible for humans to detect (Murphy, 2012).

In his paper "Some Studies in Machine Learning Using the Game of Checkers" from 1959, Arthur Samuel describes his pioneering work in the field of ML, specifically the development of a computer program that could play and improve at the game of checkers. This program, called the "Checker-Playing Program", was notable for its use of a self-learning algorithm that could improve its performance over time through experience and trial-and-error.

Samuel's approach was based on the idea of reinforcement learning, which involves training a machine learning algorithm through a series of rewards and punishments. In the case of the Checker-Playing Program, the algorithm was rewarded for making good moves and punished for making bad ones. Over time, the program learned to associate certain board configurations with positive or negative outcomes and used this knowledge to make better decisions in future games.

2.2.2 Deep Learning

Deep learning (DL) is a subset of machine learning that is based on artificial neural networks with multiple layers of abstraction, including an input layer, one or more hidden layers, and an output layer. Each layer consists of multiple nodes or neurons, which are connected to the nodes in the previous and next layers (Russell & Norvig, 2010).

DL, much like Machine Learning, uses development of algorithms and patterns to allow systems to learn and develop an understanding of unknown and previously untaught situations and data. During the training phase, the weights of the connections between the neurons are adjusted to minimize the error between the predicted output and the actual output. DL is currently at the forefront of numerous applications in various fields such as image and speech recognition, natural language processing, and robotics.

One of the key features of DL is the ability to learn end-to-end mappings between the input data and the output, which can be used for tasks such image classification and object detection. Another important feature is the ability to automatically learn hierarchical representations of data, which can be used to extract meaningful features and reduce the dimensionality of the data (Russell & Norvig, 2010).

A notable application of DL is the field of Natural Language Processing (NLP), where DL models have achieved impressive results on tasks such as machine translation, sentiment analysis, and text classification. Kim (2014) presents a DL architecture called the convolutional neural network (CNN) that can effectively model the local and global features of text data. The author demonstrates the effectiveness of this approach on several NLP tasks, achieving state-of-the-art results on sentiment analysis and topic classification benchmarks.

2.2.3 Teaching the AI

There are many ways of teaching an AI, and adding data points for it to analyze and base decisions on. However, thus far there is no definitive “right” way of doing it. There are different schools of thought and action that are more effective depending on the usage of the AI and amount of data it’s going to be basing decisions on.

Supervised learning is one of the main strategies to teach AI, it entails the training of algorithms using data, where input-output pairs serve as examples. The algorithm learns patterns, enabling it to generalize and make predictions on new, unknown data. Our main concern with supervised learning is the possibility of bias in the training data, which can result in discriminatory behavior in the AI system (Goodfellow, Courville & Bengio, 2016). This can be particularly problematic when the training data is not representative of the entire population, such as when data is collected from a specific demographic group. An example of this would be Amazon's hiring bias debacle, which reportedly made it difficult for any non male applicant to be considered, since most of the AIs training data came from male submitted resumes (Dastin, 2018). Additionally, there is a risk of overfitting, which occurs when the AI system becomes too specialized in the training data and is unable to generalize to new situations (Goodfellow, Courville & Bengio, 2016).

Reinforcement learning (RL) is a type of machine learning in which an agent learns to interact with an environment in order to maximize a cumulative reward signal (Sutton, 2018). Unlike supervised learning, in which the agent is given labeled data to learn from, and unsupervised learning, in which the agent must find structure in unlabeled data, RL requires the agent to learn from feedback signals received after taking actions in an environment. The goal of RL is to learn a policy, which is a mapping from states to actions, that maximizes the expected cumulative reward over time. One of the key advantages of RL is that it can learn optimal behavior in complex and uncertain environments, where it may not be feasible to calculate all possible states and actions (Sutton, 2018).

2.3 Defining bias

The definition of bias is “the action of supporting or opposing a particular person or thing in an unfair way, because of allowing personal opinions to influence your judgment” (Cambridge Dictionary, n.d.). In the context of recruitment, bias is an important factor to acknowledge, as the decisions made may affect people's career and livelihood. Bias itself has a lot of subsections. This thesis will focus on the three subsections of racial bias, gender bias and sexuality bias.

2.3.1 Gender bias

Gender bias refers to the systematic favoritism or discrimination towards individuals based on their gender during the recruitment process. It can manifest in various ways, such as biased language in job descriptions, gendered expectations, or unconscious biases during candidate evaluation. (Brescoll, Uhlmann & Newman, 2013). Gender bias may lead to underrepresentation of certain genders in specific roles or industries, perpetuating gender inequalities.

2.3.2 Racial bias

Racial bias involves the unequal treatment of individuals based on their race or ethnicity during recruitment. It can occur through various means, including discriminatory hiring practices, biased resume screening, or subjective evaluations influenced by stereotypes. Racial bias not only limits opportunities for individuals from marginalized racial or ethnic groups but also perpetuates social inequalities and hampers diversity and inclusion efforts within organizations (Pager et al., 2009).

2.3.3 Sexuality bias

Sexuality bias pertains to prejudice or discrimination based on individuals' sexual orientation. It involves negative attitudes, stereotypes, or discriminatory actions towards individuals who identify as lesbian, gay, bisexual, or other sexual orientations (Pager et al., 2009). Sexuality bias can manifest in various forms, including harassment, exclusion, stigma, and denial of rights or privileges. It perpetuates social inequalities and denies individuals the ability to express their authentic selves. Overcoming sexuality bias involves promoting acceptance, understanding, and equal treatment for individuals across the spectrum of sexual orientations.

2.4 Bias within recruitment

The recruitment process plays a crucial role in shaping the composition and diversity of the workforce in various industries, including the IT sector. However, it has been widely acknowledged that bias within recruitment can hinder equal opportunities and perpetuate inequalities. This section aims to explore the topic of bias within recruitment, through the three main lenses mentioned above.

Bias depending on gender remains a significant concern in recruitment practices. Otugo et al. (2020) highlight the prevalence of potential bias against women and underrepresented individuals, and the increase of this bias as a result of modern recruitment techniques like virtual interviews. Things such as non-verbal cues and unconscious biases, can unintentionally disadvantage women and contribute to the perpetuation of gender disparities in the IT sector (Otugo et al., 2020). Furthermore, research has historically shown that women get paid less for similar work in international organizations. Filmer, King & Van de Walle

(2005) also show potential discrimination in pay, explaining that “The difference in mean current salaries between men and women (holding nationality group constant) is 14 percent, almost half of which cannot be attributed to differences in observed characteristics such as age, education, and previous work experience.” (Filmer, King & Van de Walle, 2020, p.19). Reducing this bias is possible however, through actions such as using gender-neutral language in job advertisements, and having women in leadership positions (Chatterjee & Shenoy, 2023). Women also have a greater perceived prevalence of gender bias in other areas such as education (García-González, Forcen & Jimenez-Sanchez, 2019), which may also apply to areas such as recruitment.

Similarly, racial bias in hiring processes is an area of critical examination. Edward and Park (1999) shed light on the existence of racial bias within recruitment, emphasizing the impact it has on equitable access to opportunities. Their research highlights the need to address biases in the evaluation and selection of candidates to promote a more inclusive and diverse IT workforce. The pay within international organizations depends on your country of operation. People from developed countries that have access to bank loans earn 5% more money than people from countries without this access (Filmer, King & Van de Walle, 2005). It should be noted that in the case of this pay difference, it is more likely to be due to structural changes than differences in characteristics according to the authors. Filmer, King & Van de Walle (2005). also mention that the hiring process is shifting, where it may become less biased as more women and people from developing countries enter the workforce.

In the context of sexuality bias, Everly, Unzueta, and Shih (2015) conducted a study that sheds light on the biases faced by gay and lesbian job applicants. Their research findings revealed intriguing insights into how sexual orientation impacts perceived hireability. The study found that men perceived gay and lesbian job applicants as less hirable compared to heterosexual job applicants. On the other hand, women tended to perceive gay and lesbian applicants as more hirable than heterosexual applicants. These findings underscore the presence of bias based on sexual orientation and its potential influence on hiring decisions, and shows that a variety of people may be needed within an organization in order to reduce hiring bias, as predicted by Filmer, King & Van de Walle. (2005).

2.5 Bias within AI

The development of artificial intelligence (AI) systems entails multiple stages that can contribute to the emergence of bias. This section will delve into the details of how bias can manifest within AI, focusing on data collection and preprocessing, algorithmic design, and system implementation.

2.5.1 Data collection and preprocessing

Bias may originate during the data collection phase, wherein training data can mirror historical societal inequities, giving rise to biased patterns (Jussupow et al. 2021). Biases in data can stem from several sources, such as human-generated labels**,** data sampling techniques, or data collection processes influenced by societal stereotypes or prejudices (Barocas & Selbst, 2016). Moreover, it is important to acknowledge that the human factor plays a crucial role in the data collection process. Human biases can inadvertently seep into the data collection phase, whether through subjective labeling, biased sampling techniques, or the influence of prior beliefs and stereotypes (Friedler et al., 2018). These human biases can introduce skewed or unrepresentative data, which in turn may lead to biased outcomes when training AI algorithms. These biases can be perpetuated when training AI models, subsequently leading to biased outcomes in recruitment decisions. Addressing this challenge necessitates meticulous attention to data collection methodologies, ensuring the inclusion of diverse and representative data samples while minimizing the impact of biased annotations. While AI algorithms are intended to operate without human biases, the reality is that the design of these algorithms involves human choices and subjective decision-making. Human biases can manifest in algorithmic design through the selection of features, the formulation of optimization objectives, or the choice of evaluation metrics (Kusner et al., 2017). Therefore, it becomes crucial to critically evaluate the design choices and ensure that human biases are not inadvertently embedded into the algorithms.

2.5.2 Algorithmic Design

Bias can also manifest within AI through choices made during algorithmic design. Biases may arise from the mathematical modeling and optimization objectives of AI algorithms. For instance, if a model is trained to minimize prediction errors without explicitly considering fairness criteria, it can inadvertently learn and perpetuate discriminatory patterns present in the data (Bolukbasi et al., 2016). Algorithmic biases can manifest in various forms, such as differential error rates across different demographic groups or biased predictions based on protected attributes. Tackling algorithmic bias calls for the development of fairness-aware algorithms that incorporate fairness metrics and constraints into the optimization process.

2.5.3 System implementation

Further, bias within AI can develop during the implementation phase. System design choices, including feature selection, performance metrics, and decision thresholds, can introduce or amplify biases in AI applications. Biased outcomes may arise if certain features disproportionately influence the decision-making process or if performance metrics do not account for fairness considerations (Caliskan et al., 2017). Moreover, human biases can inadvertently permeate AI systems through the subjective interpretation and encoding of decision rules. As the implementation of AI systems involves human agency and decision-making, which can introduce or perpetuate biases. Factors such as the interpretation of AI outputs, the setting of decision thresholds, or the consideration of contextual factors are influenced by human judgment. These subjective decisions can inadvertently amplify or reinforce existing biases in the AI system (Kleinberg et al., 2017). It is imperative to have mechanisms in place to address potential biases during the system implementation phase and ensure that human interventions are guided by fairness considerations. Mitigating bias at the implementation level necessitates careful system design, integrating fairness-aware considerations and ensuring transparency in decision-making processes. An example of AI system bias is Microsoft Tay, a chatbot launched in 2016. Designed to interact with users on social media, Tay swiftly adopted offensive and biased language due to exposure to inflammatory content, underscoring the potential pitfalls of inadequate data filtering and insufficient oversight in AI implementation (Vincent, 2016).

2.5.4 Trust and AI

Users' interactions with AI systems span a wide spectrum, from seeking recommendations for everyday choices like driving routes or content preferences to relying on AI algorithms for complex decision-making in fields such as finance and healthcare. However, alongside the growing reliance on AI, there emerges a complex interplay between trust and fear, especially when it comes to the potential for AI to make incorrect decisions (Cui, 2023).

Trust is a cornerstone of human relationships, and its extension to interactions with technology, including AI, is crucial for establishing confidence in the system's capabilities. Users tend to trust AI systems that consistently provide accurate and relevant recommendations, thereby simplifying their decision-making processes. This trust is often reinforced by AI's ability to process vast amounts of data, recognize patterns, and offer insights that human judgment might miss. When AI consistently delivers positive outcomes, users feel more comfortable relying on its advice, resulting in a sense of partnership between humans and machines (Cui, 2023).

However, this evolving trust in AI is not without its apprehensions. One significant fear is the possibility of AI making incorrect decisions, which can have varying degrees of impact depending on the context. Too much trust in a system can sometimes lead to conformity. Liel & Zalmanson (2020) found that clearly visible mistakes were trusted by users, because of previously high accuracy. While AI's processing capabilities are formidable, it lacks the depth of human intuition and contextual understanding. Users may fear that AI might fail to account for nuanced factors or unexpected variables, leading to misguided recommendations or decisions that do not align with human values.This fear of AI making wrong decisions can be exacerbated by a lack of transparency and explainability (Förster et al. 2020). If AI operates as a "black box," (Schmid, 2023) providing recommendations without offering insight into how those recommendations were generated, users may become wary of blindly following AI advice. The inability to understand the reasoning behind AI's suggestions can hinder users' willingness to trust its decisions, as well as their ability to criticize it (Förster et al., 2020).

Developers and designers are working towards enhancing the transparency of AI systems. An example of this would be to build their decision-making processes to have tools like interpretability tools and visualization dashboards to make them more comprehensible to users (Raji, 2019). By providing explanations for AI-generated recommendations, users can better evaluate the credibility of the advice and gain insights into the factors considered by the AI. Proactive steps, such as auditing training data for bias, employing diverse teams during development, and continually monitoring AI outcomes, can ensure fairness and rectify discrepancies that erode trust. Encouraging users to critically assess AI recommendations, cross-referencing them with their own judgment, can mitigate the risk of blind conformity (Raji, 2019).

2.6 Thematic Overview

The thematic overview serves as a summary of the topics and aspects of the literature review. Furthermore, it serves as a guide for creating the questionnaire used for data collection.

Table 2.1: Thematic Overview

Theme	**Sub-theme**	**Literature**
The Recruitment Process	Stages of the recruitment process	(Cambridge Dictionary, N.D.), (Cook, 2009), (Indeed, 2023), (Kethley & Terpstra, 2005)
Understanding AI, Machine Learning, and Algorithms	What is AI	(Nilsson, 1998), (Russell & Norvig, 2010)
	Machine learning	(Jordan & Mitchell, 2015), (Murphy, 2012), (Samuel, 1959)
	Deep learning	(Kim, 2014), (Russell & Norvig, 2010)
	Teaching the AI	(Dastin, 2018), (Goodfellow, Courville & Bengio, 2016), (Sutton, 2018)
Defining Bias	What is bias	(Cambridge Dictionary, n.d.)
	Gender bias	(Brescoll, Uhlmann & Newman, 2013)
	Racial bias	(Pager et al., 2009)

	Sexuality bias	(Pager et al., 2009)
Bias within Recruitment	Gender Bias	(Chatterjee & Shenoy, 2023), (Filmer, King & Van de Walle, 2005), (García-González, Forcen & Jimenez-Sanches, 2019), (Otugo et al., 2020)
	Racial Bias	(Edward & Park, 1999), (Filmer, King & Van de Walle, 2005)
	Sexuality Bias	(Everly, Unzueta & Shih, 2015), (Filmer, King & Van de Walle., 2005)
Bias within AI	Data collection and preprocessing	(Barocas & Selbst, 2016), (Friedler et al., 2018), (Jussupow et al., 2021), (Kusner et al., 2017)
	Algorithmic Design	(Bolukbasi et al., 2016)
	System implementation	(Caliskan et al., 2017), (Kleinberg et al., 2017), (Vincent, 2016)
	Trust and AI	(Cui, 2023), (Förster et al, 2020), (Raji, 2023), (Schmid, 2023)

3 Research Methodology

3.1 Search Strategy

As mentioned in 1.1, the main purpose of this study is to answer the research question: "How do recruiters in the Swedish IT sector perceive the effectiveness of AI tools in mitigating bias within the recruitment process?". As such, researching the three areas of Bias, Recruitment and Artificial Intelligence is crucial, as well as their connections and any gaps in literature.

Using databases and tools such as Google Scholar, LUBSearch, EBSCOHost Research Databases and SpringerLink, the literature was selected based on articles and books from several different journals, after analyzing their impact scores and relevancy in their respective subjects. ResearchGate was also used in order to view articles, but only in cases in which the article was confirmed to be published in a high impact score journal and peer reviewed and was not accessible to us directly through that journal due to fees or inability to download the document.

Because the literature covers three main areas: Bias, AI and Recruitment, several different sets of keywords were used, which include:

Table 3. 1: Keywords used for literature review

Keywords	In combination with:
AI Artificial Intelligence Bias Gender Sexuality Racial Discrimination Machine Learning Deep learning Recruitment Selection Fairness Deep Learning Preprocessing Algorithmic design System implementation	within in definition

3.2 Method of Data Collection

In order to gain a higher number of respondents, a self-administered questionnaire was chosen as the data collection method. The questionnaire aims to focus on a mix of factual data and opinionated data, through the combination of open and closed questions.

The benefit of a questionnaire is the fact that the answers are easier to analyze as researchers, and they are easy for a respondent to complete compared to for example a full interview. In addition, the ability for the questionnaire to be self-administered lowers the risk for changes in answer due to the behavior of the researchers or administrators (Oates, 2006).

Closed questions allow for direct analysis to be done on the amount of answers provided, while open questions may provide different viewpoints to the researches previously not thought of, and may uncover a problem in the area of recruitment that may not be very explored yet. Closed questions aim to provide a pattern in answers, where open ended ones may bring guidance for potential issues or benefits that the administrator had not previously thought of (Oates, 2006).

However, it is also important to consider disadvantages with choosing questionnaires as a data collection method. Firstly, respondents are not able to ask researchers questions directly, which may lead to frustration of the respondent, or alternatively an inability to answer the questions themselves. Additionally, the truthfulness of the answers cannot be checked by the researchers (Oates, 2006). As such, it is vital when creating the questionnaire to give clear and concise questions that are easy to answer for people working in the field of recruitment, and to be able to verify the answers in other means, especially when making a Web-Based questionnaire.

According to Oates (2006), the disadvantages of questionnaires include the fact the researchers cannot provide clarity of the subject, and respondents often cannot motivate their answers. In order to combat this, each part ends with optional open-ended questions, where respondents can clarify and motivate parts of their answers. The questionnaire is divided into an initial information page describing the purpose and intent of the research and the contact information to the researchers, as well as three other sections.

Section 1 consists of control questions. The main objective is to gather information about the respondent, while not collecting any compromising personal information such as their name or where they work in accordance with The General Data Protection Regulation (GDPR) (European Commision, n.d.). The first two questions verify that the respondent is qualified to answer the questionnaire. Any respondents not answering “Yes” to both of these are excluded from data analysis.

Section 2 deals with the issue of bias within recruitment, where respondents answer questions related to bias within the current recruitment process, such as the prevalence of different types of bias, or the change of it during their time as a recruiter. An optional, open-ended question is provided at the end of the section where the respondent is able to describe the main challenges they face in addressing bias as part of recruiting.

Section 3 addresses the possibility of implementing AI as part of the recruitment process. The section begins with a short text, explaining terminology used. Respondents provide information about whether they use AI in recruiting, and the potential help and harm AI tools could cause within the industry according to them. Two optional, open-ended questions are provided at the end, where respondents can describe the ways AI has helped them in recruiting and any incidents where AI would cause problems respectively.

A pilot study was sent out to 5 respondents that were not part of the target group, with the main purpose of identifying any potential errors in the layout and structure of the questionnaire, and any potential issues in clarity. According to Oates (2006), a questionnaire needs to be both well-made and attractive to increase response rate. This pretest was not made to evaluate the questions, but instead the structure of the questionnaire itself, with the purpose of making it easy to follow along and increase the number of respondents as a result. Following the completion of the pilot study, the questionnaire was distributed to the previously decided target group, where errors were corrected in the structure and one question that was not deemed clear and purposeful towards the study was removed.

All questions can be found in “Appendix A – Questions”.

3.3 Sample Criteria, Respondents, Questionnaire

The target group for the research was defined as “Recruiters who work for an IT company, or for a recruitment firm specializing in IT positions, for the Swedish job market”. As such, control questions were set up in order to assure the reliability of the respondents’ answers. The respondents were selected based on these criteria, and sent the questionnaire, either via contact on LinkedIn, or through E-mail.

3.4 Ethical Aspects

According to Oates (2006), several ethical aspects must be considered when conducting research, including the right not to participate, the right to withdrawal, the right to give informed consent, the right to anonymity and the right to confidentiality.

The right to not participate was considered through the initial information page of the questionnaire, stating “participation is entirely voluntary”. The purpose and intent of the research has also been made clear here, as well as what the data will be used for. The contact information of the researchers was provided, where respondents can contact the researchers in order to have their answers deleted, or alternatively have their answers be made confidential.

Respondents remain anonymous through the questions asked, where name and affiliation is not provided as part of the answers. Contacting specific respondents through the channels mentioned in 3.3 ensures that the data collected remains reliable while minimizing the potential risks of gathering sensitive data.

3.5 Validity and reliability

Validity and reliability are both important matters when conducting research. Validity is divided into two parts, internal and external validity.

Internal validity measures the ability of the research to be relevant to the subject it is trying to measure (Jacobsen, 2017). In order to meet this criteria and make the research relevant, a comprehensive theoretical overview was made, which was then used to create the table of questions used in the questionnaire for data collection, shown in “Appendix I - Questions”. It is very important in quantitative research to make sure that every question asked in a questionnaire is well formulated and relevant to the overall purpose, as the data collected is only as good as the questions asked (Oates, 2006).

External validity measures the ability of the research to be generalized and therefore able to be applied to a bigger audience or population (Jacobsen, 2017). The benefits in a quantitative method lies in the larger amount of responses, which helps contribute to the external validity of the research. This makes it more likely that the answers represent a reliable portion of the target group.

Reliability is the ability of the research to be trusted (Jacobsen, 2017). Simply collecting a large amount of data is not helpful in itself, as the target group must still be well defined, preventing answers from unqualified respondents that could negatively affect the reliability of the results. As such, the respondents were chosen according to the method in 3.3 with the purpose of balancing the size of the target group with the amount of respondents.

3.6 Reflection on Methodology

The decided method and approach was deemed appropriate for the scope and purpose of the research, providing useful insight into the opinion on AI use and implementation in recruitment processes. However, it should be noted that the amount of responses gathered means a lower statistical significance than what may be needed to achieve conclusive results. The research method produced results that were deemed to be satisfactory to the scope of the bachelor thesis. However, the method could be improved through an increased number of respondents which could increase external validity of the results, opening up the possibility of further data analysis, such as the Pearson correlation coefficient.

Alternatively, a mixed methods approach could be used in larger scale research. This method would consist of questionnaire data listed in 3.2 and analysis in part 3.3 to gather the general attitude and thoughts towards AI use in recruitment. Interviews would also be conducted with experts in implementing AI systems to get insight into the implementation of AI tools from a development standpoint. The benefits of such a method would be to both gather an overall picture of the use of AI tools within recruitment on a larger scale, as well as getting a deeper insight into the technical implementations of AI tools and the possibility to debias them.

Furthermore, calculating the Pearson correlation coefficient between results and giving the reliability an exact numerical value could be an improvement on the legibility of the results, giving a numerical correlation value that shows the pattern between the different questions analyzed in “Appendix D - Response Analysis”.

4 Results

“This chapter presents the findings from the analysis of the collected data through questionnaires”. Each of the first three parts starts with presenting each relevant question, with a description of the responses. The final part summarizes the results in table form.

4.1 Control Questions

The first question of our questionnaire was one of two control questions which aimed to ensure validity of the answers. Out of the respondents, three people (15%) answered that they do not work within the recruitment sector. Thus three of the respondents' answers needed to be excluded from the data.

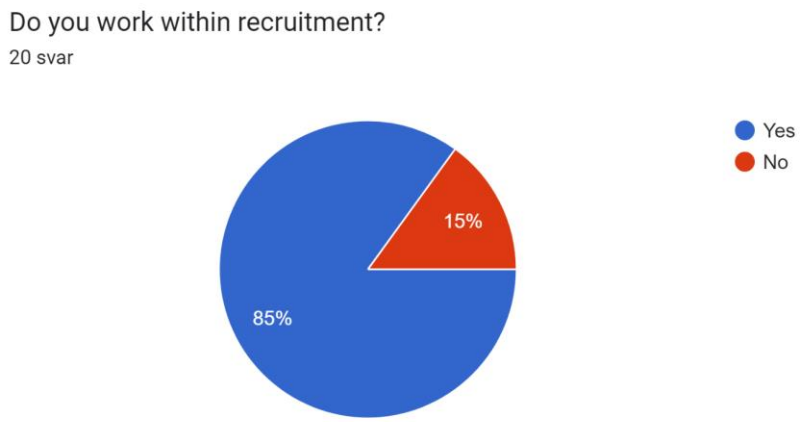

Figure 4.1: Do you work within recruitment?

Out of 20 total respondents, three (15%) answered that they do not work for an IT company or a recruitment agency within the IT sector. Thus there were additional respondents’ answers which had to be excluded from the data.

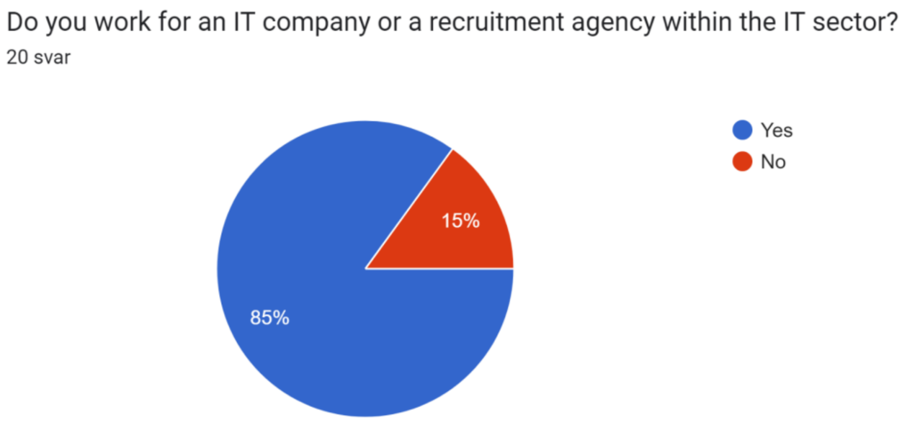

Figure 4.2: Do you work for an IT company or a recruitment agency within the IT sector?

A total of four respondents had to be removed, because of failed control questions. One respondent answered negatively only on the question “Do you work within recruitment?”.. One respondent answered negatively only on the question “Do you work for an IT company or a recruitment agency within the IT sector?”. Two respondents answered negatively on both questions, totalling 4 ineligible respondents. Therefore the presented data for the following questions will include only 16 respondents.

Out of 16 total respondents, a majority (68.8%) answered that they would describe themselves as interested in or knowledgeable of AI within recruitment. Five respondents (31.3%) answered that they do not.

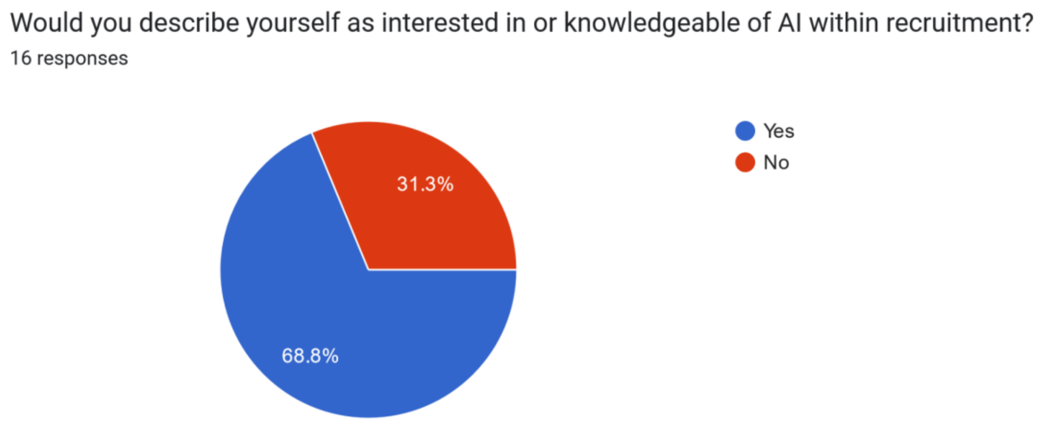

Figure 4.3: Would you describe yourself as interested in or knowledgeable of AI within recruitment?

4.2 Bias within Recruitment

The question "What is your assessment of the prevalence of bias in the current recruitment process?" is formatted in the form of a scale question, where respondents answered between option 1 (Negligible Bias) and 5 (Significant bias). The results show a neutral trend, with a most recurring value, or mode, of 3 (Average bias), and a mean of 2.875 (Slightly below average bias)

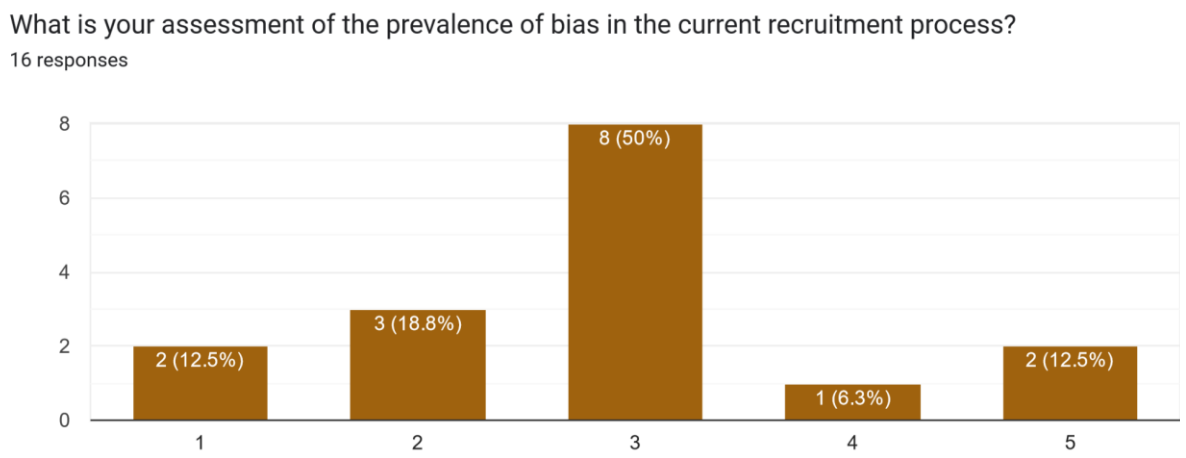

Figure 4.4: What is your assessment of the prevalence of bias in the current recruitment process?

The question “How much do you think personal preferences and subjective judgements of human recruiters contribute to bias in the selection of candidates?” is formatted in the form of a scale question, where respondents answered between option 1 (No Role) and 5 (A very big role). The results show a trend towards a bigger role in the selection, with a mode of 4 (6 respondents) and a mean of 3.5 (In Between an average and big role).

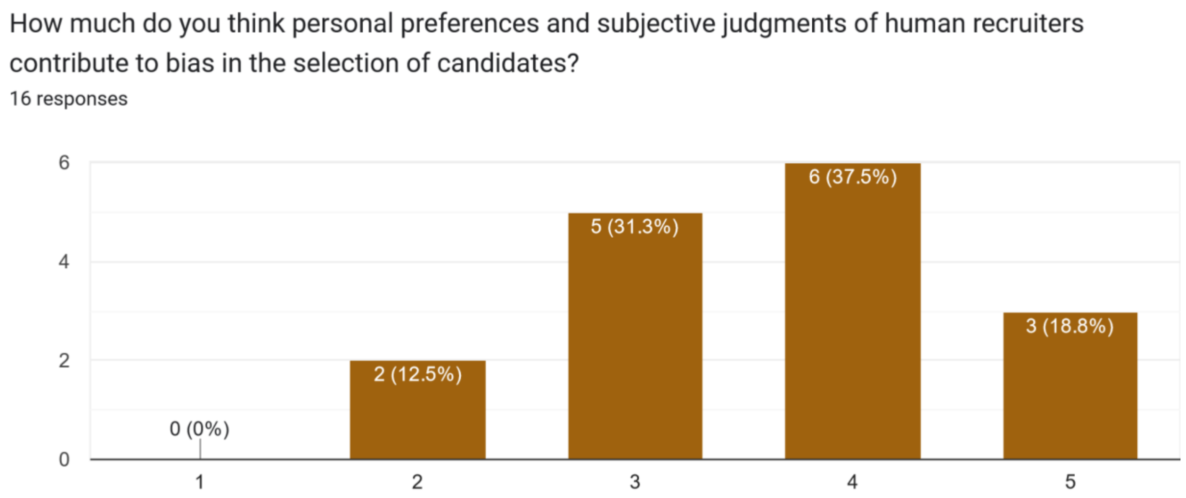

Figure 4.5: How much do you think personal preferences and subjective judgements of human recruiters contribute to bias in the selection of candidates?

Four respondents to the open-ended question highlighted cultural differences as the main challenge when facing bias in the recruitment process. One respondent listed the need to make sure that everyone else in the process is unbiased, while one respondent listed denial and low level of self awareness as one of the main challenges.

The question “How confident are you in your ability to identify and address bias in the recruiting process?” is formatted in the form of a scale question, where respondents answered between option 1 (Not confident) and 5 (Extremely confident). The results show a positive trend, indicating that recruiters on average are very confident in their ability, with a mode of 4 and a mean of 4.

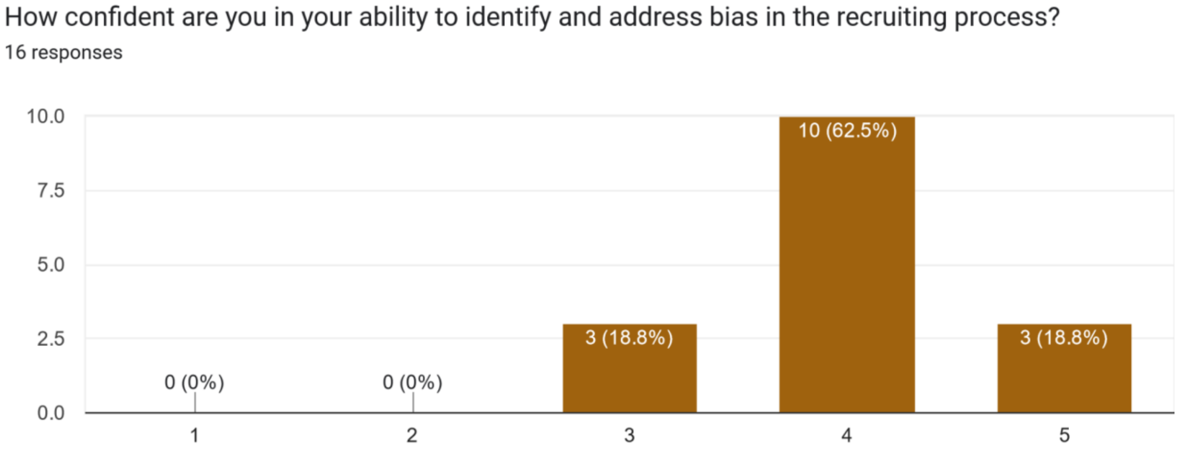

Figure 4.5: How confident are you in your ability to identify and address bias in the recruiting process?

The question “How common has racial bias been within the industry during your time as a recruiter?” is formatted in the form of a scale question, where respondents answered between option 1 (Non-existent) and 5 (Extremely common). The results are mixed, with responses leaning slightly in the negative (not common). The most recurring value is 2 (not common) and the mean is 2.75.

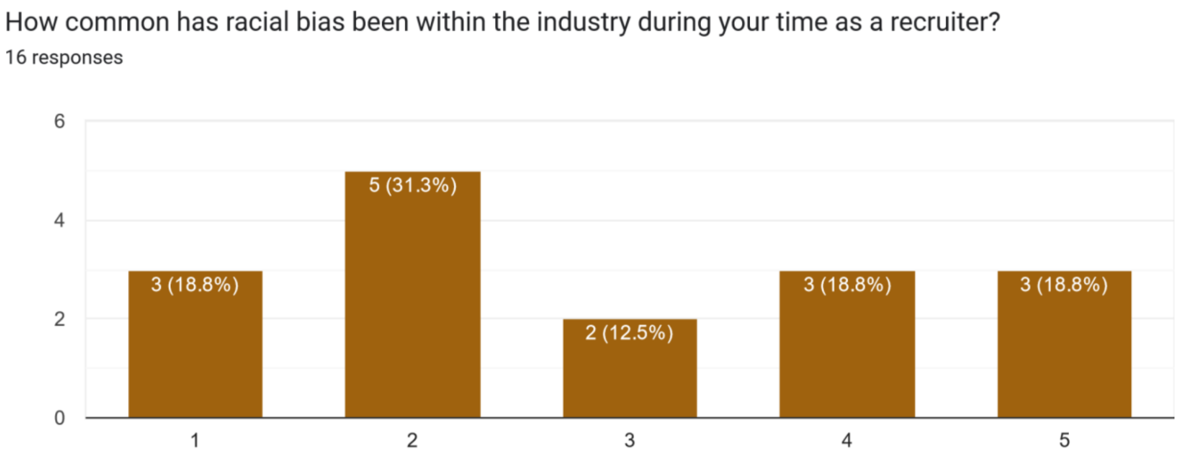

Figure 4.6: How common has racial bias been within the industry during your time as a recruiter?

The question “What changes have you observed in the prevalence of racial bias during your time as a recruiter?” is formatted in the form of a scale question, where respondents answered between option 1 (A lot less bias) and 5 (A lot more bias). The results show a trend towards less bias over time. The most recurring value is 2 (less bias) and the mean is 2.3125.

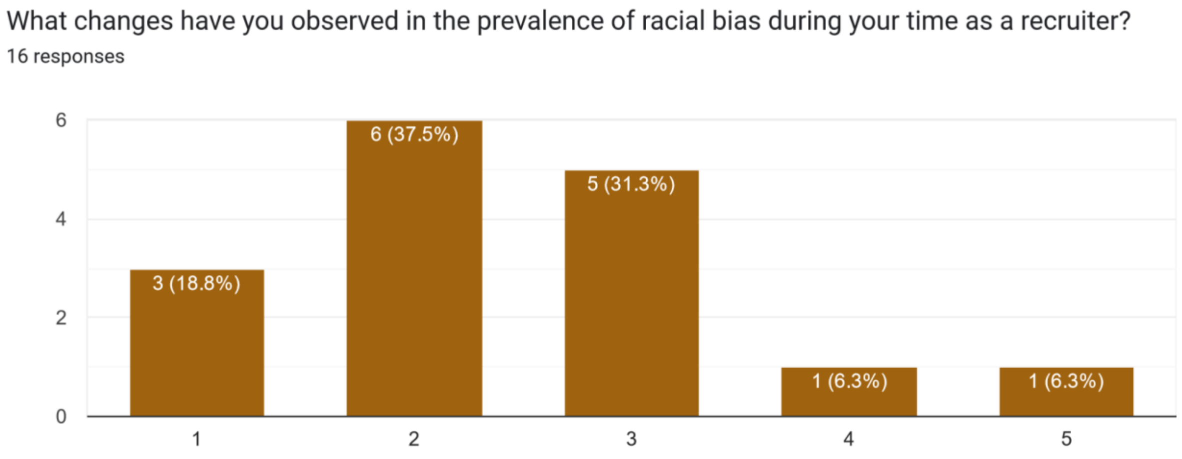

Figure 4.7: What changes have you observed in the prevalence of racial bias during your time as a recruiter?

The question “How common has gender bias been within the industry during your time as a recruiter?” is formatted in the form of a scale question, where respondents answered between option 1 (Non-existent) and 5 (Extremely common). The results show a trend towards less bias. The most recurring value is 2 (not common) and the mean is 2.

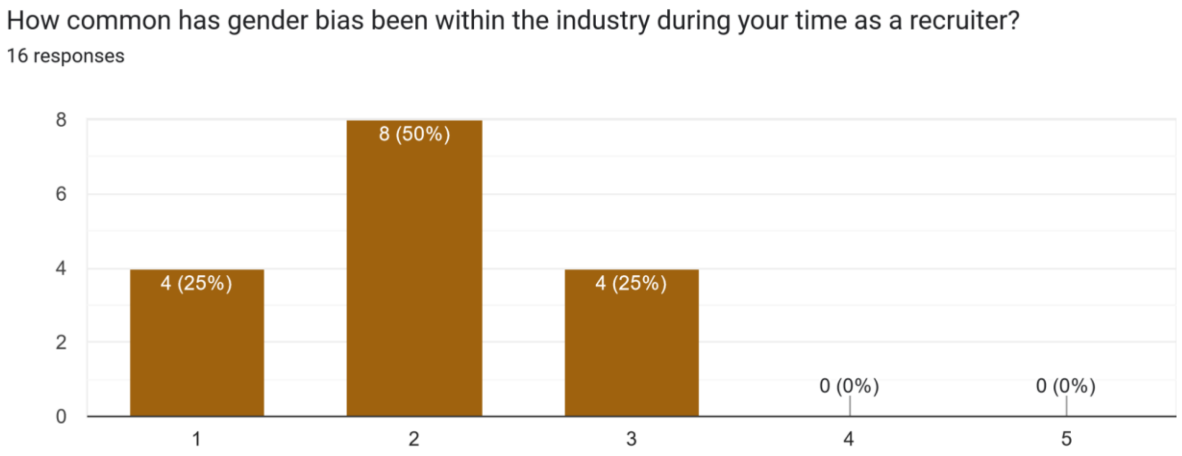

Figure 4.8: How common has gender bias been within the industry during your time as a recruiter?

The question “What changes have you observed in the prevalence of gender bias during your time as a recruiter?” is formatted in the form of a scale question, where respondents answered between option 1 (A lot less bias) and 5 (A lot more bias). The results are mixed, leaning slightly towards the lower end. The most common answers are 1 (A lot less bias), 2 (less bias) and 3, (No change). The mean is 2.5, showing an overall slight decrease in the prevalence of gender bias over time.

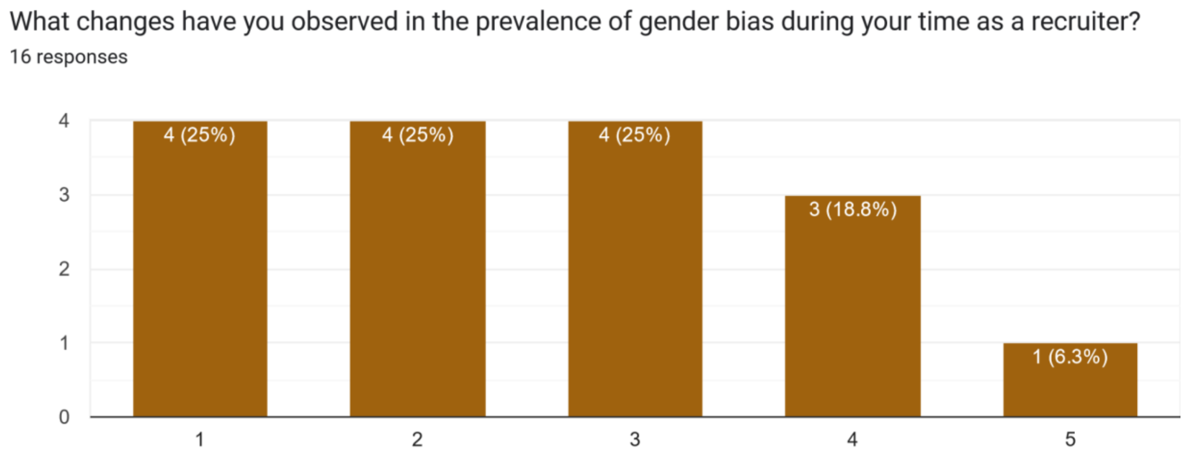

Figure 4.9: What changes have you observed in the prevalence of gender bias during your time as a recruiter?

The question “How common has sexuality bias been within the industry during your time as a recruiter?” is formatted in the form of a scale question, where respondents answered between option 1 (Non-existent) and 5 (Extremely common). The results show a trend towards less bias. The most recurring value is 1 (non existent) and the mean is 1.9375.

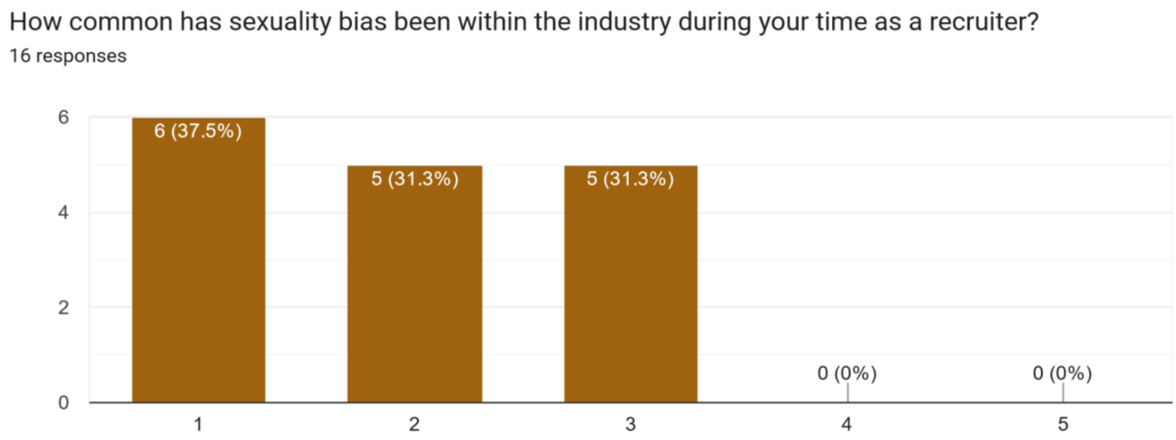

Figure 4.10: How common has sexuality bias been within the industry during your time as a recruiter?

The question “What changes have you observed in the prevalence of sexuality bias during your time as a recruiter?” is formatted in the form of a scale question, where respondents answered between option 1 (A lot less bias) and 5 (A lot more bias). The results are mixed, with a slight trend towards less bias. The most occurring value, the mode, is 3 (no change) and the mean is 2.5625, showing an average decrease in the prevalence of sexuality bias over time.

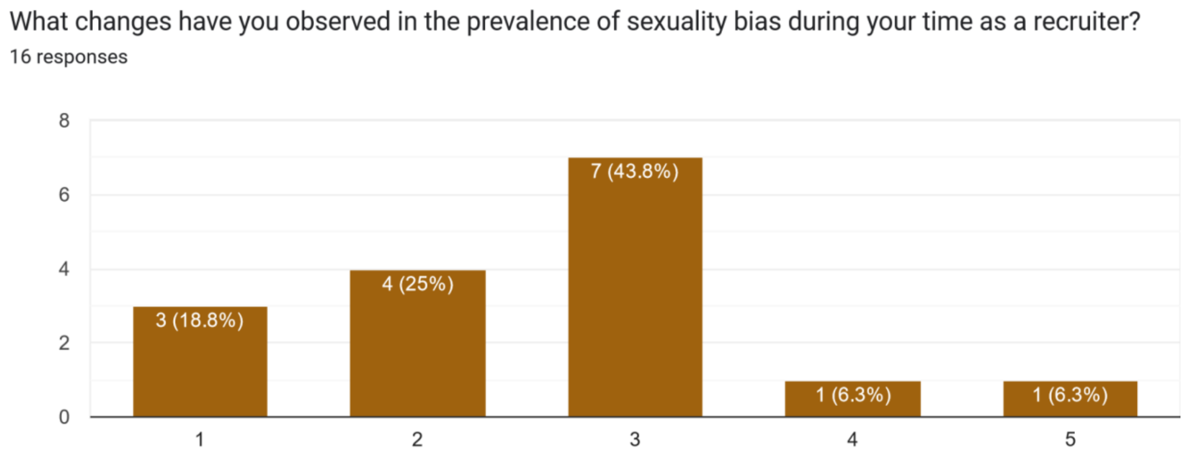

Figure 4.11: What changes have you observed in the prevalence of sexuality bias during your time as recruiter?

4.3 Bias within AI

Our first question in the section “AI-based recruitment and its effect“ has to do with the respondents hands-on experience with AI in their current companies recruitment process. The question was “As far as you know, does your company use AI in any part of the recruitment process?”. The possible answers was Yes(37.5%), No(56.3%), Not comfortable answering(0%) and Don’t know(6.3%).

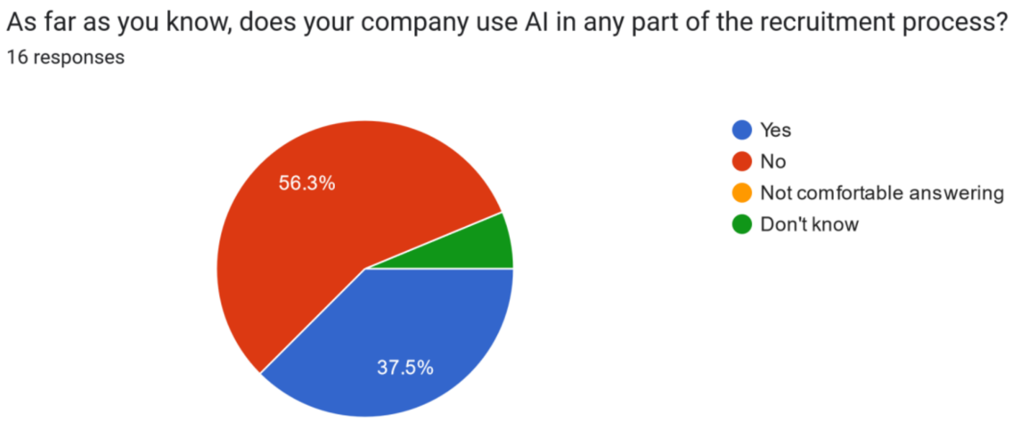

Figure 4.12: As far as you know, does your company use AI in any part of the recruitment process?

The question "How effective do you perceive current AI technologies to be in mitigating bias in recruitment processes?" has the format of a scale from "Ineffective" (1) to "Extremely effective"(5). Results here lean towards being decently effective, however it's clear there's room for improvement, since no respondents picked the most positive option. The most votes were placed on option 4(effective) and the average opinion for this question is therefore 2.8125.

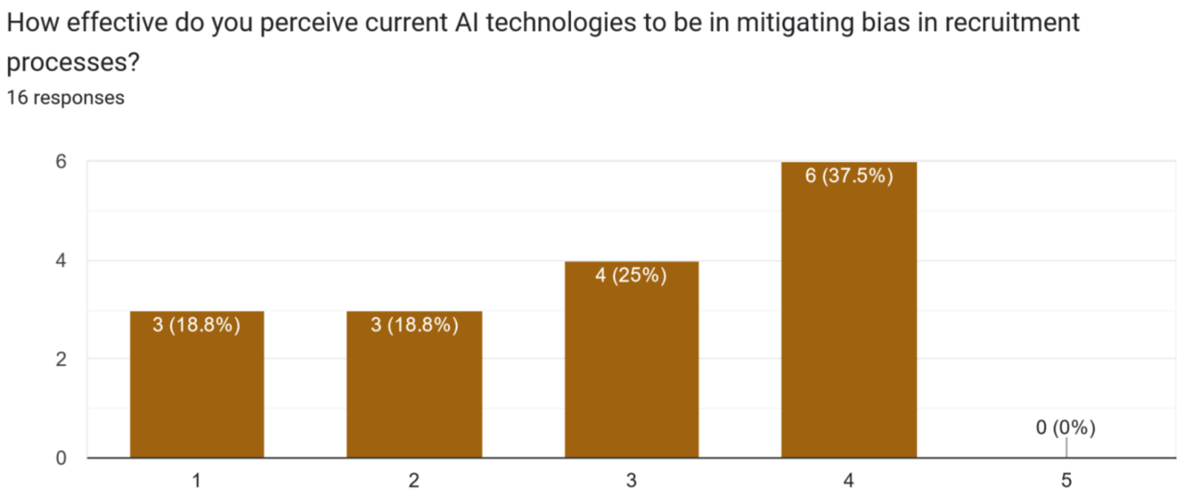

Figure 4.13: How effective do you perceive current AI technologies to be in mitigating bias in recruitment processes?

The question "To what extent are you afraid of an AI making a selection decision based on unfair bias?" has the format of a scale from "Not afraid" (1) to "Extremely afraid" (5). The results show a small to decent amount of fear, but again nothing in the extreme. The most votes were placed on option 3(afraid) and the average opinion for this question is therefore 2.4375.

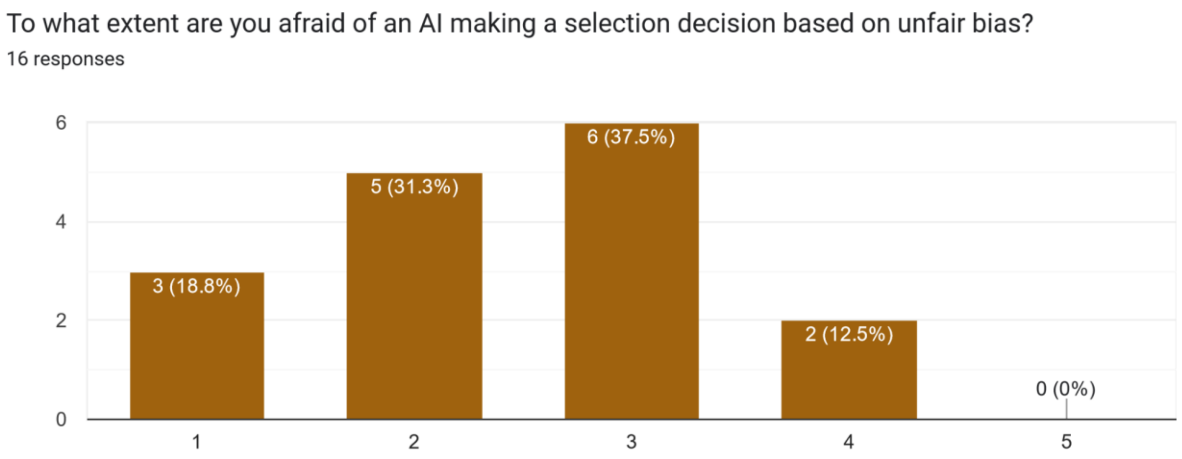

Figure 4.14: To what extent are you afraid of an AI making a selection decision based on unfair bias?

The question "How big of an issue do you believe the risk for bias to be as a result of insufficient data input?" has the format of a scale from "Not an issue"(1) to "An extremely big issue"(5). All our respondents agree it is an issue, however it's not an extremely big issue to most. The most votes were placed on option 4(a big issue) and the average opinion for this question is 3.25.

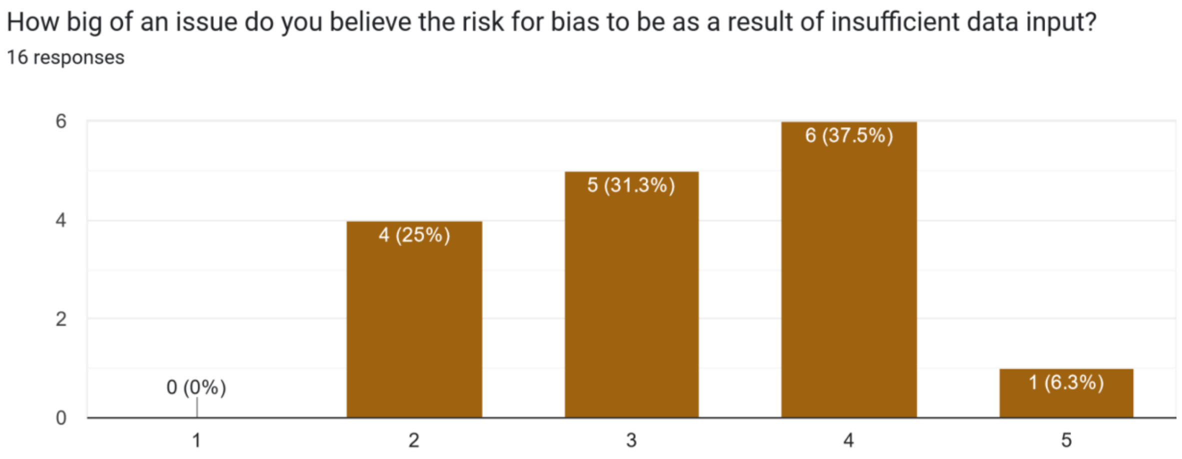

Figure 4.15: How big of an issue do you believe the risk for bias to be as a result of insufficient data input?

Two responses from the provided open questions highlighted the importance of bad data input as a problem within AI technologies being used in recruitment processes, specifically due to the use of same-company data or same-sector data that defines "successes" traditionally. For the full answers to the open-ended questions, refer to "Appendix III - Open Ended Questions".

The question "How big of an issue do you believe the risk for bias to be as a result of a bad algorithmic design?" has the format of a scale from "Not an issue"(1) to "An extremely big issue"(5). Not all our respondents agree it is an issue, however the overall option is leaning towards this being quite a big issue. The most votes were placed on option 4(a big issue) and the average opinion for this question is 3.5625.

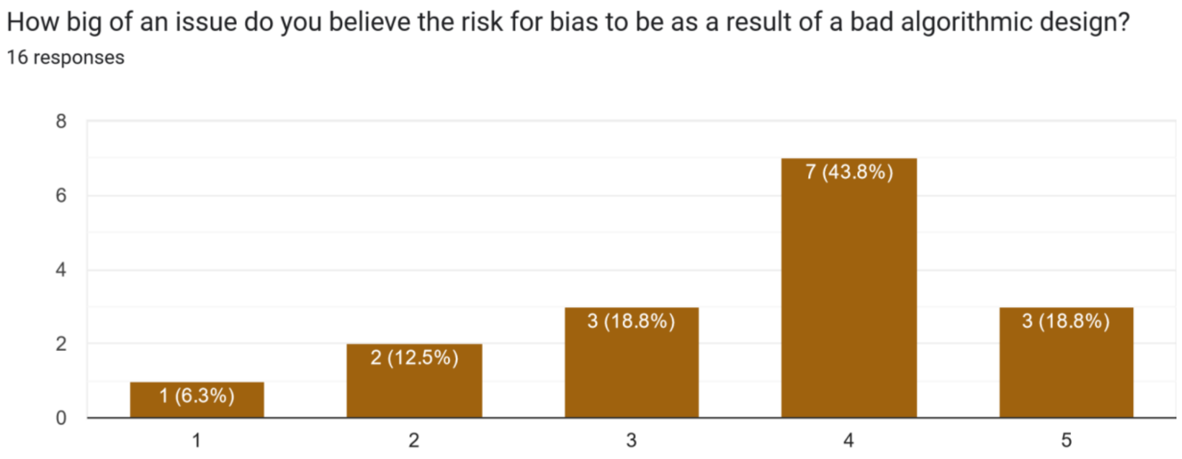

Figure 4.16: How big of an issue do you believe the risk for bias to be as a result of a bad algorithmic design?

The question "How big of an issue do you believe the risk for bias to be as a result of a bad system implementation?" has the format of a scale from "Not an issue"(1) to "An extremely big issue"(5). This question is a decisive one, since most of the votes land in the higher tiers of the question. The most votes were placed on option 4 and 5(an extremely big issue), and the average opinion for this question is 3.75.

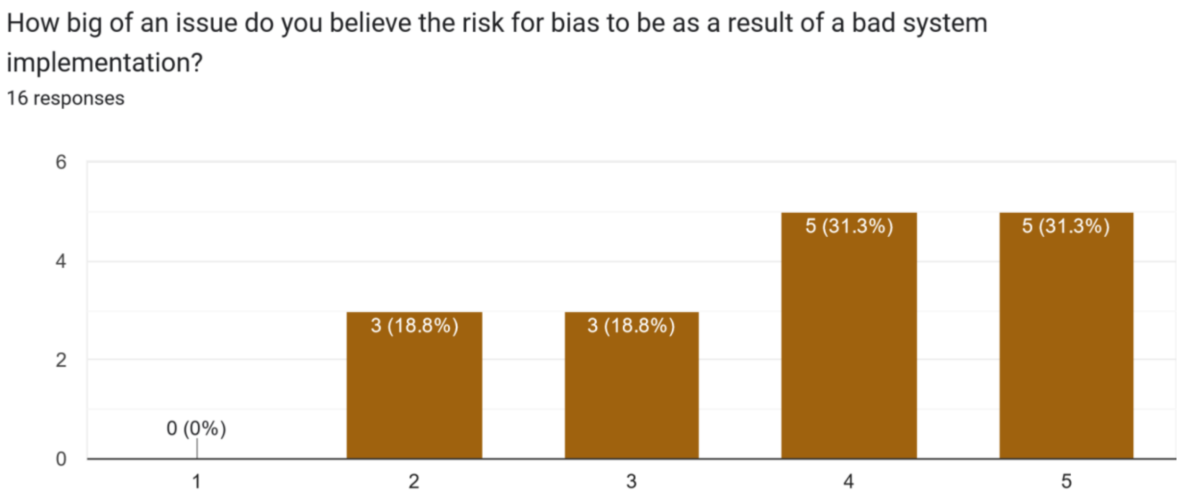

Figure 4.17: How big of an issue do you believe the risk for bias to be as a result of a bad system implementation?

The question “How big of a role do you believe human intervention should play in the decision-making process when using AI for recruitment?” has the format of a scale from “No role”(1) to “A very big role”(5). This question is also a decisive one, since most of the votes land on option 4 and 5 of the question. The most votes were placed on option 4(A big role), and the average opinion for this question is 3.8125.

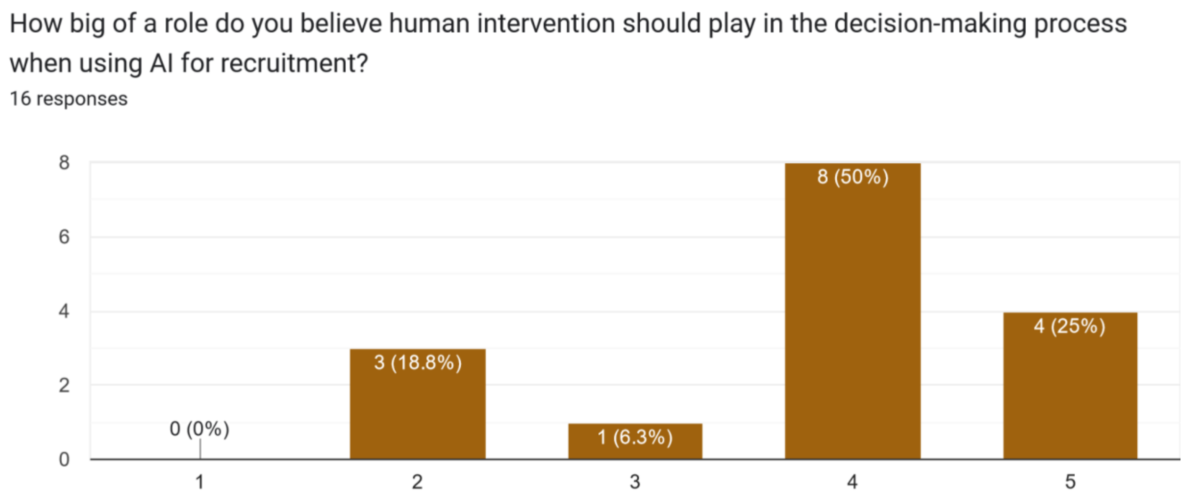

Figure 4.18: How big of a role do you believe human intervention should play in the decision-making process when using AI for recruitment?

4.4 Summary of results

Below is a table that summarizes the findings found in the graphs, Areas marked with a * have been analyzed according to the tables in “Appendix D - Response Analysis“.

Table 4. 1: Summary of Results

Area	**Main Findings**
Bias within Recruitment	● Bias is still prevalent in the current recruitment process, and is overall quite common. ● According to respondents, human subjective judgements and unconscious bias contribute to a large portion of bias during selection, despite respondents being very confident in their own ability to identify and address bias. ● All three types of bias appear to still exist in the selection process, and are perceived to be common. ● The most common type of bias in selection according to the study is racial bias, followed by gender bias and sexuality bias. ● All three types of bias have gotten less common over time. Gender bias has seen the largest decline, followed by racial bias and sexuality bias.
Bias within AI	● A minority (37.5%) of respondents both use and are aware of using AI as part of their recruitment process. ● AI technologies are not perceived to be very effective at their current level in recruitment. ● A majority of respondents feel that human intervention should play a major role in the decision-making process when using AI for recruitment.
Response Analysis*	● Female respondents perceived gender bias as more prevalent than male respondents, and have seen a lesser decline in its prevalence over time. ● Female respondents perceived sexuality bias as more prevalent than male respondents, and have seen a lesser decline in its prevalence over time. ● Male respondents perceived racial bias as more prevalent than female respondents, but female respondents have seen a lesser decline in its prevalence over time. ● Years of experience with recruitment does not seem to affect the confidence in respondents ability to identify and address bias.

5 Discussion

The analysis of the collected empirical material provided us with valuable perceptions and attitudes towards AI-based recruitment, specifically concerning bias mitigation. The discussion below addresses our findings and compares them with previous literature in the context of bias in recruitment processes and AI implementation.

In this essay we investigated how recruiters in the Swedish IT sector perceive the effectiveness of AI tools in mitigating bias within the recruitment process. Participants' positive view of AI's role in bias mitigation resonates with the growing interest in utilizing technology to reduce subjectivity in decision-making (Cui, 2023). However, intriguingly, women perceive a slightly higher degree of gender bias in recruitment than other genders, consistent with previous studies on the prevalence of gender bias in other contexts, such as education (García-González, Forcen & Jimenez-Sanchez, 2019). This suggests that women may be more attuned to gender-related biases and underscores the need to diversify the workforce, as shown by Chatterjee & Shenoy (2023). Furthermore, the differential perception of bias prevalence between genders raises important questions about the impact of societal and cultural factors on individuals' perceptions. The finding that female respondents have observed a lesser decline in the prevalence of gender and sexuality bias over time highlights the complex nature of bias reduction efforts. It underscores the challenges in achieving significant and sustained progress in combating biases that have deep-seated roots in societal norms and structures.

Interestingly, male respondents perceived racial bias as more prevalent than female respondents did. This finding suggests that different genders may perceive and prioritize various forms of bias based on their unique experiences and perspectives. However, the observation that female respondents have seen a lesser decline in the prevalence of racial bias over time indicates the interconnections between different types of biases and the intricacies of systemic inequalities. The trend of declining bias prevalence over time reflects a positive shift in recruitment practices. Notably, overall gender bias is perceived to have the most significant decrease, followed by racial and sexuality bias. This positive trajectory resonates with the efforts of organizations and institutions to enhance diversity and inclusivity in their hiring practices, as speculated by Filmer, King & Van de Walle (2005).

The finding that the relationship between years of experience in recruitment and respondents ability to identify and address bias raises important questions about the relationship between experience and bias awareness. Traditionally, one might assume that individuals with more experience in the recruitment process would possess a heightened ability to recognize and mitigate bias. One plausible explanation for this finding is that the recruitment process itself is embedded with bias that becomes normalized over time. Moreover, the results invite us to examine the extent to which training and education programs within the recruitment field effectively address bias. It is possible that training initiatives focused on bias awareness and mitigation may not be reaching experienced recruiters or may not be designed to adapt to their changing needs. Ensuring that such programs remain relevant and engaging for all levels of experience could be crucial to fostering bias awareness and effective intervention.

Another noteworthy aspect that emerges from the data is the paradoxical relationship between respondents' confidence in identifying bias and the persistence of human subjective judgments leading to bias. Despite respondents' strong belief in their ability to recognize and mitigate bias, the reported prevalence of unconscious bias suggests a misalignment between self-assessment and actual outcomes, which could negatively influence AI recruitment algorithms according to Kleinberg et al. (2017).

The absence of significant fear regarding AI making unfair decisions suggests a baseline level of trust in AI systems. Yet, this trust might be influenced by participants' limited exposure to AI decision-making outcomes(Cui, 2023). Moreover, concerns about algorithmic design and system implementation taking precedence over insufficient data input underscore the importance of focusing on comprehensive system development. The unanimous agreement on the importance of human intervention and oversight in AI-based recruitment aligns with previous studies on human-machine collaboration(Förster, 2020). This recognition of human judgment and ethical considerations signifies a conscious effort to avoid over-reliance on technology and to ensure accountability.

One of our findings, that only a minority of respondents (37.5%) both use and are aware of using AI in their recruitment process does not reflect previous findings, as presented on ResumeBuilder (2023). However, this may indicate that the current state of adoption of AI tools within the Swedish IT-industry is lagging behind. Moreover, the big difference in adoption of AI could be explained by the magnitude of the company, or what work is being done. This suggests that while AI has gained traction, there is still a significant portion of organizations that have not embraced AI technologies in their recruitment practices. The relatively low awareness and adoption of AI, combined with the perceived ineffectiveness of current AI applications, indicate that the industry is at a crossroads in leveraging AI for recruitment. While AI technologies hold promise for streamlining processes and mitigating bias, it is evident that they are not yet meeting expectations in terms of efficacy. Whatever the reason, from the trust of the AI to how efficient it might be considered, it’s clear AI still has a way to go, before it’s common practice in IT recruitment.

6 Conclusion

The purpose of this thesis was to better understand how recruiters perceive AI tools and their ability to mitigate bias as part of the recruitment in Swedish IT companies.

Our results indicate that recruiters do not trust people to always make the right decisions, and to be unbiased. This could possibly be due to multifaceted differences in upbringing and underlying unconscious human bias in the recruiter, combined with lack of self-awareness and difficulty in identifying personal bias. They do trust AI, to not make unfair decisions, to a slightly higher degree, but only with a high amount of human supervision. This aligns with previous research showing that AI tools may not work optimally unless supervised by human users. Taking all these statements into consideration, we believe AI, in its current form, is not a sustainable or effective option for debiasing the recruitment process in the IT industry. Further progress in algorithmic design and implementation needs to be made before a trustworthy AI is created. Examples include easy ways of having oversight, as well as a transparent way for companies to gain insight into why and how decisions are made.

Our findings suggest that issues like trust and acceptance of the tools would arise in wider usage of the currently available AI. However, the perceived issues in the current recruitment processes circulate around human error in the form of personal preferences and subjective human judgment. Having cultural and societal values will always be a hindrance when trying to be fully neutral, and that is something AI can be devoid of. It’s therefore our belief that with proper oversight and support, AI can be a very useful tool to diminish bias in the recruitment processes in IT companies.

6.1 Limitations

During the writing process, limitations and weaknesses of the research method have become clear. A quantitative method has its positives in the amount of data it can gather, and the numerical analysis possible due to directly comparable answers in the form of a number. However, due to the amount of responses being relatively low, we believe our findings may not be seen as indisputable answers. However, larger scale investigations may yield such results.

6.2 Future Work

The field of AI and bias within recruitment presents several possibilities for future research. An improved understanding of the users of AI tools is needed, but other areas also need to be understood in order to effectivize its use. This includes understanding the human factors that go into developing AI tools for use in recruitment, and the direct comparison of different tools and their usefulness, with and without human collaboration. While this thesis focuses on how recruiters see AI tools and their use in the industry, direct comparison between an AI-based recruitment and a human-based recruitment is essential to fully understand its effect on bias.

Appendix A - Questions

Below shows the initial information page and all questions used in the respective parts mentioned in chapter 3.3 - Data collection and analysis, as well as their formats. Question formats have been derived from Oates (2006).

Questionnaire: "AI and the Future of IT Recruitment"

We are two bachelor students from Lund University, writing our thesis on Design of Information Systems.

The research examines the possibility of using Artificial Intelligence tools within recruitment and its potential effects on recruitment bias.
This questionnaire aims to provide context of bias within the current recruitment system as well as take a look at any potential pros and cons with implementing this new technology within IT recruitment. The survey takes 10-15 minutes to complete.

All data collected will be treated with strict confidentiality. The information you provide will be used solely for the purpose of this research study.
Your responses will be anonymized and aggregated to ensure that individual participants cannot be identified in any published results or reports. Rest assured that your privacy and confidentiality will be maintained throughout the study.

Participation in this study is entirely voluntary. You have the right to withdraw from the study at any time without penalty or prejudice. By proceeding with the questionnaire, you are indicating your consent to participate. If you decide to withdraw, simply close the browser or exit the survey.

If you have any questions or concerns about this study or the data collection process, please feel free to contact the researchers:

John Arvidsson
Email: jo3086ar-s@lu.se
Linkedin: https://www.linkedin.com/in/john-arvidsson-08a769179/

Ludvig Eslind
Email: lu1445es-s@lu.se
Linkedin: https://www.linkedin.com/in/ludvig-eslind-910176210/

By clicking the "Next" button below, you confirm that you have read and understood the above information and voluntarily consent to participate in this research study

Figure A 1: Questionnaire information page

Table A 1: Basic questions, format included

Question:	Question format:
Do you work within recruitment?	Yes/no
Do you work for an IT company or a recruitment agency within the IT sector?	Yes/no
Would you describe yourself as interested in or knowledgeable to AI within recruitment?	Yes/no
What is your gender identity?	Multiple Choice (Male, Female, Other, Prefer not to indicate)
What is your age?	Multiple Choice (Below 25, 25 to 35, 36 to 45, above 46)
Please choose the option which best reflects your current position.	Multiple Choice (Senior/Executive Management, Middle Management/Supervisor, Support Staff/Administrative, Professional/Individual Contributor)
Please indicate the amount of time you have worked at your current organization.	Multiple Choice (Less than 5 years, 5-9 years, 10-19 years, Longer than 20 years)

Table A 2: Bias in recruitment questions, format included

Question:	Type of Question:
What is your assessment of the prevalence of bias in the current recruitment process?	Scale question (1-5)
How much do you think personal preferences and subjective judgments of human recruiters contribute to bias in the selection of candidates?	Scale question (1-5)
How confident are you in your ability to identify and address bias in the recruiting process?	Scale question (1-5)
How common has racial bias been within the industry during your time as a recruiter?	Scale question (1-5)
What changes have you observed in the prevalence of racial bias during your time as a recruiter?	Scale question (1-5)
How common has gender bias been within the industry during your time as a recruiter?	Scale question (1-5)
What changes have you observed in the prevalence of gender bias during your time as a recruiter?	Scale question (1-5)
How common has sexuality bias been within the industry during your time as a recruiter?	Scale question (1-5)
What changes have you observed in the prevalence of sexuality bias during your time as a recruiter?	Scale question (1-5)
What are the main challenges you face in addressing bias in the recruitment process?	Scale question (1-5)

Table A 3: AI-based recruitment and its effect questions, format included

Question:	Type of Question:
As far as you know, does your company use AI in any part of the recruitment process?	Scale question (1-5)
How effective do you perceive current AI technologies to be in mitigating bias in recruitment processes?	Scale question (1-5)
To what extent are you afraid of an AI making a selection decision based on unfair bias?	Scale question (1-5)
How big of an issue do you believe the risk for bias to be as a result of insufficient data input?	Scale question (1-5)
How big of an issue do you believe the risk for bias to be as a result of a bad algorithmic design?	Scale question (1-5)
How big of an issue do you believe the risk for bias to be as a result of a bad system implementation?	Scale question (1-5)
How big of a role do you believe human intervention should play in the decision-making process when using AI for recruitment?	Scale question (1-5)
Have you found any specific ways AI has assisted or helped you work within recruitment?	Open Question (Optional)
Have you encountered any specific challenges or limitations in implementing AI-based recruitment solutions in your organization? If so, please describe.	Open Question (Optional)

Appendix B - Graphs

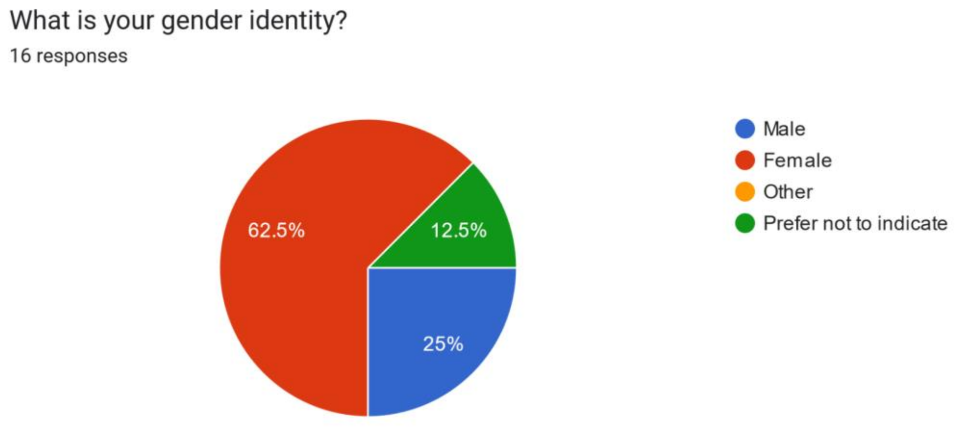

Figure B 1: What is your gender identity?

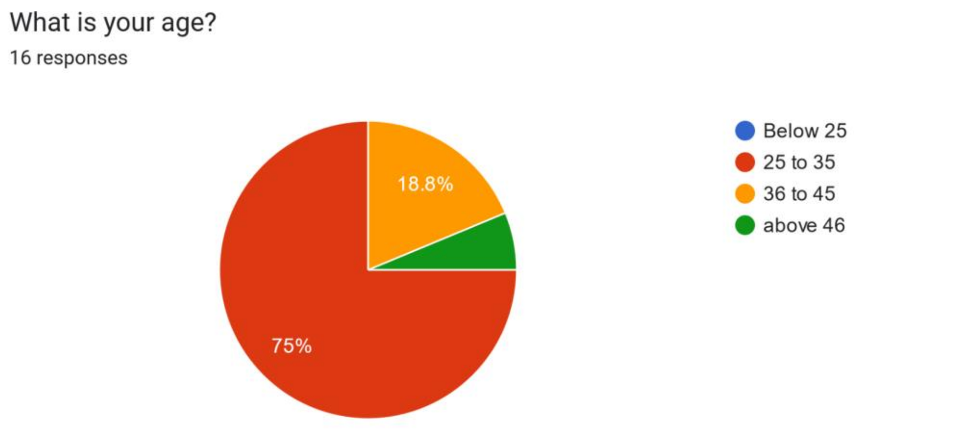

Figure B 2: What is your age?

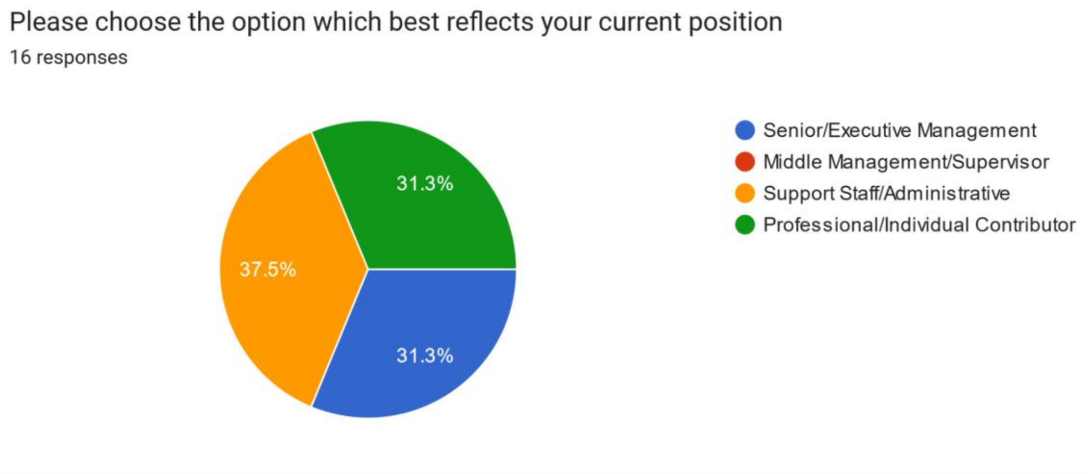

Figure B 3: Please choose from the option which best reflects your current position

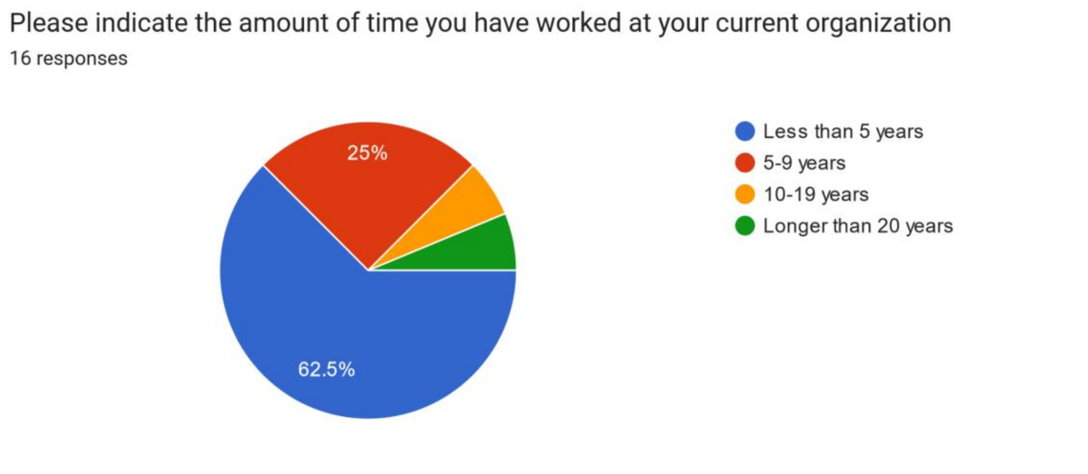

Figure B 4: Please indicate the amount of time you have worked at your current organization

Appendix C - Open Ended Questions and Answers

What are the main challenges you face in addressing bias in the recruitment process?

7 svar

Denial, low levels of self awareness and cultural roots

cultural challenges

Stereotypes are sometimes hard to get away from, if we truly scrutinise ourselves i believe there's always bias to be found. So getting away from unconscious bias i would say is the hardest part.

Language requirements

cultural differences and language

Prejudice

To make sure everyone else in the process of hiring someone is free from bias

Figure C 1: What are the main challenge you face in the recruitment process?

Have you found any specific ways AI has assisted or helped you work within recruitment?

7 responses

Recommendation alg, matching candidates and positions can be effective.

not yet

Not this far. Perhaps in screening candidates.

no

previewing potential candidates

No

Yes, sometimes I take the persons profile from linked in and ask AI to summarize it. Or I use AI to make conclusions from my summary from the interview. I also ask AI a lot of questions and advice. Somtimes specifik questions

Figure C 2: Have you found any specific ways AI has assisted or helped you work within recruitment?

Have you encountered any specific challenges or limitations in implementing AI-based recruitment solutions in your organization? If so, please describe.

6 responses

No

It will learn from the company historical data and definition of success, that already has a lot of bias.

No.

Data might already be skewed since most likely it will be taken from the company or sector itself

We can not implement it in the process since we work mostly with headhunting, meaning we search for a specific candidate and it´s really hard to even recive answers from the ones we want. But we do implement it in the process like I said above; to analyze interviews or ask AI for advice etc

Figure C 3: Have you encountered any specific challenges or limitations in implementing AI-based recruitment solutions in your organisation? If so, please describe.

Appendix D - Response Analysis

The below table shows cross analysis between respondents' responses to the question “Please indicate the amount of time you have worked at your current organization” and (“How big of a role do you believe human intervention should play in the decision-making process when using AI for recruitment?”).

Table D 1: Relation between years of experience in recruiting on significance of human intervention when using AI for recruitment

	Years of Experience			
	<5	**5-9**	**10-19**	**20+**
Amount of human intervention (1-5):	5	4	2	2
	4	5		
	4	5		
	3	4		
	4			
	2			
	5			
	4			
	4			
	4			
Total:	3.9	18	2	2
Mean:	39	4.5	2	2

Cross analysis between respondents response to the question “Please indicate the amount of time you have worked at your current organization” and “How confident are you in your ability to identify and address bias in the recruiting process?”.

Table D 2: Relation between years of experience in recruiting on confidence in ability to identify and address bias in the recruitment process.

	Years of Experience			
	<5	**5-9**	**10-19**	**20+**
	5	3	4	4
	4	4		
	4	4		
	5	5		

Confidence in identifying bias (1-5)	4			
	4			
	4			
	3			
	3			
	4			
Total:	40	16	4	4
Mean:	4	4	4	4

Cross analysis between respondents' response to the question "Please indicate the amount of time you have worked at your current organization" and "To what extent are you afraid of an AI making a selection decision based on unfair bias?".

Table D 3: Relation between years of experience in recruiting on fair of AI making unfair decisions.

	Years of Experience			
	<5	**5-9**	**10-19**	**20+**
Fear of AI making unfair decisions (1-5)	5	4	2	2
	4	5		
	4	5		
	3	4		
	4			
	2			
	5			
	4			
	4			
	4			
Total:	39	18	2	2
Mean:	3.9	4.5	2	2

Cross analysis between respondents' responses to the question “Please indicate your gender”, and the prevalence of gender bias, as well as its change over time. Question type indicates the question answered, Prevalence is the question “How common has Gender bias been within the industry during your time as a recruiter?” and Change is the question “What changes have you observed in the prevalence of gender bias during your time as a recruiter?”. Respondents choosing “prefer not to indicate” as their gender were not analyzed in this question.

Table D 4: Relation between gender and perceived prevalence and change of gender bias.

	Prevalence (Gender Bias)		Change (Gender Bias)	
Gender:	**Female**	**Male**	**Female**	**Male**
Answer (1-5):	3	1	2	1
	3	3	3	4
	2	2	3	1
	2	1	4	2
	1		1	
	3		4	
	2		2	
	2		3	
	2		1	
	2		3	
Total:	22	7	26	8
Mean:	2.2	1.75	2.6	2

Cross analysis between respondents' responses to the question “Please indicate your gender”, and the prevalence of racial bias, as well as its change over time. Question type indicates the question answered, Prevalence is the question “How common has racial bias been within the industry during your time as a recruiter?”. Change is the question (“What changes have you observed in the prevalence of racial bias during your time as a recruiter?”). Respondents choosing “prefer not to indicate” as their gender were not analyzed in this question.

Table D 5: Relation between gender and perceived prevalence and change of racial bias.

	Prevalence (Racial Bias)		Change (Racial Bias)	
Gender:	**Female**	**Male**	**Female**	**Male**
	2	1	2	1

	5	5	3	4
	3	5	3	1
Answer (1-5):	4	3	2	2
	1		1	
	4		2	
	4		3	
	2		3	
	2		2	
	2		3	
Total:	29	14	24	8
Mean:	2.9	3.5	2.4	2

Cross analysis between respondents' responses to the question “Please indicate your gender”, and the prevalence of sexuality bias, as well as its change over time. Question type indicates the question answered, Prevalence is the question “How common has sexuality bias been within the industry during your time as a recruiter?” and Change is the question (“What changes have you observed in the prevalence of sexuality bias during your time as a recruiter?”). Respondents choosing “prefer not to indicate” as their gender were not analyzed in this question.

Table D 6: Relation between gender and perceived prevalence and change of sexuality bias.

	Prevalence (Sexuality Bias)		Change (Sexuality Bias)	
Gender:	**Female**	**Male**	**Female**	**Male**
Answer (1-5):	2	1	2	1
	3	3	3	3
	2	1	2	2
	3	2	3	3
	1		1	
	3		4	
	1		1	
	1		3	
	3		3	
	2		3	
Total:	21	7	25	9
Mean:	2.1	1.75	2.5	2.25

References

1 in 4 companies have already replaced workers with ChatGPT. (2023, February 27). Resume Builder, Available online: https://www.resumebuilder.com/1-in-4-companies-have-already-replaced-workers-with-chatgpt/ [Accessed 30 March 2023]

Barocas, S. & Selbst, A. D. (2016). Big Data's Disparate Impact, 104 California Law Review 671 (2016), Available Online: http://dx.doi.org/10.2139/ssrn.2477899 [Acessed 18 August 2023]

Bolukbasi, T., Chang, K., Zou, J., Venkatesh, S. & Kalai, A. (2016). Man is to Computer Programmer as Woman is to Homemaker? Debiasing Word Embeddings, Available online: https://arxiv.org/abs/1607.06520 [Accessed 20 August 2023]

Bosak, J. & Sczesny, S. (2011) Gender Bias in Leader Selection? Evidence from a Hiring Simulation Study, *Sex Roles*, Available online: https://doi.org/10.1007/s11199-011-0012-7 [Accessed 11 August 2023]

Brescoll, V.L., Uhlmann, E.L. & Newman, G.E. (2013). The effects of system-justifying motives on endorsement of essentialist explanations for gender differences, *Journal of Personality and Social Psychology*, Available Online: https://psycnet.apa.org/record/2013-41938-001 [Accessed 11 August 2023]

Caliskan, A., Bryson, J. J. & Narayanan, A. (2017). Semantics derived automatically from language corpora contain human-like biases, *Science*, Available Online: https://www.science.org/doi/10.1126/science.aal4230 [Accessed 18 August 2023]

Cambridge Dictionary. (n.d.). Bias. Available Online: https://dictionary.cambridge.org/dictionary/english/bias [Accessed 17 August 2023]

Cambridge Dictionary. (n.d.). Recruitment. Available Online: https://dictionary.cambridge.org/dictionary/english/recruitment [Accessed 17 August 2023]

Chatterjee, A. & Shenoy, V. Understanding gender discrimination and biasness in the workplace: voice from women segment, *Strategic HR Review*, Available Online: https://doi-org.ludwig.lub.lu.se/10.1108/SHR-04-2023-0023 [Accessed 17 August 2023]

Cook, M. (2009). Personnel Selection: Adding Value Through People, 5th edn, John Wiley & Sons

Cui, T., Li, S., Chen, K., James, B. & Feng, L. (2023). Designing Fair AI Systems: Exploring the Interaction of Explainable AI and Task Objectivity on Users' Fairness Perception, *PACIS 2023 Proceedings.* Available online: https://aisel.aisnet.org/pacis2023/161 [Accessed 12 August 2023]

Dastin, J. (2018). Amazon scraps secret AI recruiting tool that showed bias against women, Reuters, Available online: https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G [Accessed August 11 2023]

Edward, J. & W. Park. (1999). Racial Ideology and Hiring Decisions in Silicon Valley. *Qualitative Socialogy*, Vol. 22, No. 3, pp.223-233

European Commission. (n.d). Data protection in the EU. Available Online: https://commission.europa.eu/law/law-topic/data-protection/data-protection-eu_en [Accessed 19 August 2023]

Everly, B., Unzueta, M., Shih, M. (2015). Can Being Gay Provide a Boost in the Hiring Process? Maybe If the Boss is Female, *Journal of Business and Psychology*, Available online: https://doi.org/10.1007/s10869-015-9412-y [Accessed 11 August 2023]

Filmer, D., King, E. & Van de Walle, D. (2005). Testing for pay and promotion bias in an international organization. *International Journal of Manpower*, Vol. 26 No. 5. pp 404-420.

Friedler, S. A., Scheidegger, C., Venkatasubramanian, S., Choudhery, S., Hamilton, E. P. & Roth, D. (2018). A comparative study of fairness-enhancing interventions in machine learning, *Proceedings of the Conference of Fairness, Accountability and Transparency*, Available Online: https://doi.org/10.1145/3287560.3287589 [Accessed 18 August 2023]

Förster, M. Klier, M. Kluge, K. & Sigler, I. (2020). Fostering Human Agency: A Process for the Design of User-Centric XAI Systems, *ICIS 2020 Proceedings*. Available Online: https://aisel.aisnet.org/icis2020/hci_artintel/hci_artintel/12 [Accessed 9 August 2023]

García-González, J. Forcen, P. & Jimenez-Sanchez, M. (2019). Men and women differ in their perception of gender bias in research institutions. *PLoS ONE 14(12)*, Available Online: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0225763 [Accessed 19 August 2023]

Goodfellow, I., Courville, A., & Bengio, Y. (2016). Deep learning (Vol. 1). MIT press Cambridge. Available online: https://www.deeplearningbook.org/ [Accessed 15 August 2023]

Indeed. (2023). Guide: Next Steps After You Got a Job Offer. Available Online: https://www.indeed.com/career-advice/starting-new-job/got-a-job-offer [Accessed 19 August 2023]

Jacobsen, D. I. (2017). Hur genomför man undersökningar?, Translated by S. Andersson, 2017, Studentlitteratur AB

Jordan, M. I. & Mitchell, T.M. (2015) Machine learning: Trends, perspectives, and prospects. *Science*, vol. 349, pp.255-260.

Jussupow, E. Martinez, M. Maeche, A. & Heinzl, A. (2021). Is This System Biased? - How Users React to Gender Bias in an Explainable AI System, *ICIS 2021 Proceedings.* Available online: https://aisel.aisnet.org/icis2021/hci_robot/hci_robot/11 [Accessed 9 August 2023]

Kethley, R. B. & Terpstra, D. E. (2005). An analysis of litigation associated with the use of the application form in the selection process, *Public Personnel Management*, Available Online: https://doi.org/10.1177/009102600503400407 [Accessed 12 August 2023]

Kleinberg, J., Lakkaraju, H., Leskovec, J., Ludwig, J. & Mullainathan, S. (2017) Human decisions and Machine Predictions. *The Quarterly Journal of Economics*, Available online: https://doi.org/10.1093/qje/qjx032 [Accessed 12 August 2023]

Kim, Y. (2014). Convolutional Neural Networks for Sentence Classification, *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Available online: https://aclanthology.org/D14-1181/ [Accessed 16 August 2023]

Kusner, M., Loftus, J., Russell, C. & Silva, R. (2017). Counterfactual Fairness, *31st Conference on Neural Information Processing Systems (NIPS 2017)*, Available Online: https://proceedings.neurips.cc/paper_files/paper/2017/file/a486cd07e4ac3d270571622f4f316ec5-Paper.pdf [Accessed 15 August 2023]

Liel, Y & Zalmanson, L. (2020). What if an AI Told You That 2 + 2 Is 5? Conformity to Algorithmic Recommendations, *ICIS 2020 Proceedings.* Available online: https://aisel.aisnet.org/icis2020/hci_artintel/hci_artintel/17 [Accessed 9 August 2023]

Murphy, K. (2012). Machine Learning: A Probabilistic Perspective, 2012.

Najibi, A. (2020). Racial Discrimination in Face Recognition Technology. Harvard University, Available online: https://sitn.hms.harvard.edu/flash/2020/racial-discrimination-in-face-recognition-technology/ [Accessed 11 August 2023]

Nilsson, N. J. (1998). Artificial Intelligence: A New Synthesis. Elsevier Inc. Available at: https://www.sciencedirect.com/book/9781558604674/artificial-intelligence-a-new-synthesis [Accessed 12 August 2023]

Oates, B.J. (2006). Researching Information Systems and Computing, SAGE Publications
Otugo, O., Alvarez, A., Brown, I. & Landry, A. (2020). Bias in recruitment: A focus on virtual interviews and holistic review to advance diversity, AEM Educ Train, Available Online: https://doi.org/10.1002/aet2.10661 [Accessed 11 August 2023]
P.V. Raveendra, Y. M. S. Satish & P. Singh. (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*, Available Online: https://doi.org/10.1166/jctn.2020.9086 [Accessed 11 August 2023]
Pager, D., Western, B. & Bonikowski, B. (2009). Discrimination in a Low-Wage Labor Market: A Field Experiment, *American Sociological Review 2009*, Vol. 74, pp.777-799
Raji, I. D., & Buolamwini, J. (2019). Actionable Auditing: Investigating the Impact of Publicly Naming Biased Performance Results of Commercial AI Products. *AAAI/ACM 2019 Proceedings on AI, Ethics, and Society*. Available online: https://researchr.org/publication/aies-2019 [Accessed 15 August 2023]
Sutton R, Barto, A. (2018). Reinforcement Learning: An Introduction, MIT Press (2018) second edition. Available online: https://inst.eecs.berkeley.edu/~cs188/sp20/assets/files/SuttonBartoIPRLBook2ndEd.pdf [Accessed 16 August 2023]
Russell, J. & Norvig, P. (2010). Artificial Intelligence: A Modern Approach, Pearson Education, Inc. Third edition.
Samuel, A. L. (1959). Some Studies in Machine Learning Using the Game of Checkers, *IBM Journal of Research and Development*, Vol 3. pp.210-229.
Schmid, A. & Wiesche, M. (2023). Building Trust in AI: A new understanding of the Role of Reliability, *AMCIS 2023 Proceedings*. Available online: https://aisel.aisnet.org/amcis2023/sig_aiaa/sig_aiaa/12 [Accessed 12 August 2023]
Sham, A. H., Aktas, K., Rizhinashvili, D., Kuklianov, D., Alisinanoglu, F., Ofodile, I., Ozcinar, C., & Anbarjafari, G. (2022). Ethical AI in facial expression analysis: racial bias ,*Signal, Image and Video Processing*. Available from: https://doi.org/10.1007/s11760-022-02246-8 [Accessed 30 March 2023]
Vargas, F. & Cotterell, R. (2020). Exploring the Linear Subspace Hypothesis in Gender Bias Mitigation, *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Available Online: https://aclanthology.org/2020.emnlp-main.232/ [Accessed 19 August 2023]
Vincent, J. (2016). Twitter taught Microsoft's AI chatbot to be a racist asshole in less than a day. The Verge. Available Online: https://www.theverge.com/2016/3/24/11297050/tay-microsoft-chatbot-racist [Accessed 11 August 2023]