Searching for Bias within AI-Powered Tools Used in the Recruitment Process

Implications and opportunities of predictive tools

Dimitri Reifschneider

Research Essay: AI and Society
International University of Applied Sciences

January 23, 2023

Contents

List of Figures

1. Introduction

Artificial intelligence (AI) has spread into various business sectors and workplaces as a result of technological advancements. In recent years, AI-powered tools have entered the corporate recruitment process and begun to play a significant role in decision-making. Its usage has grown rapidly in human resources (HR). In 2019, 88% of global companies were already using AI technologies in some form in HR (Brin, 2019). In fact, investments in this area are likely to grow. According to Monster (2022) 93% planned to hire in 2022 and 34% of the organizations would consider to increase the use of AI in "the event of an economic downturn ..." (Mercer, 2022, p. 3). Simultaneously, four out of ten (39%) companies want to diversify their workforce as part of their DEI (Diversity, Equity & Inclusion) objectives. But they face challenges in finding and sourcing qualified candidates, identifying matches, screening and assessing applicants (Monster, 2022). In order to counteract this, employers are relying more and more on the use of AI-powered tools. However, this solution may only appear to pursue their goals set according to the principles of diversity and equality. Indeed, AI may have negative implications for applicants.

1.1. Research Objective

The aim of this research is to determine what negative effects AI can have on equality in the recruitment process and whether it is able to make unbiased predictions. The essay defines hiring and uses its stages as a framework for further evaluation. In the end, it demonstrates possible answers to the question of how we can work toward more equity when using these tools. Following research questions should be answered and a brief overview of current AI applications should be given.

1. What are negative implications of AI in the recruitment process?
2. How can we work toward more equity and diversity when using AI in the recruitment process?

1.2. Research Outline

The first part, "Implications of AI Tools in Recruitment" (2), uses real-world examples to examine the effects of these tools. It also describes the hiring process and defines what recruitment is and how it is structured based on researched literature. The second part, "Opportunities for Equity and Diversity" (3) discusses potential recommendations and draws conclusions about how to reduce bias and improve diversity and equality.

This essay does not provide a comprehensive overview of all tools currently available on the market, nor does it go into detail about the functionalities and algorithms. It rather addresses the growing importance of existing unconscious bias in AI and shows opportunities how to address these issues. This is done by analyzing three major stages of the process: "Outreach", "Screening" and "Assessment".

Before getting into the topic, it is important to understand what hiring means. The next section defines the recruitment process, explains the relevant stages in detail and shows the role of AI. The work focuses primarily on the areas in which AI is used.

2. Implications of AI Tools in Recruitment

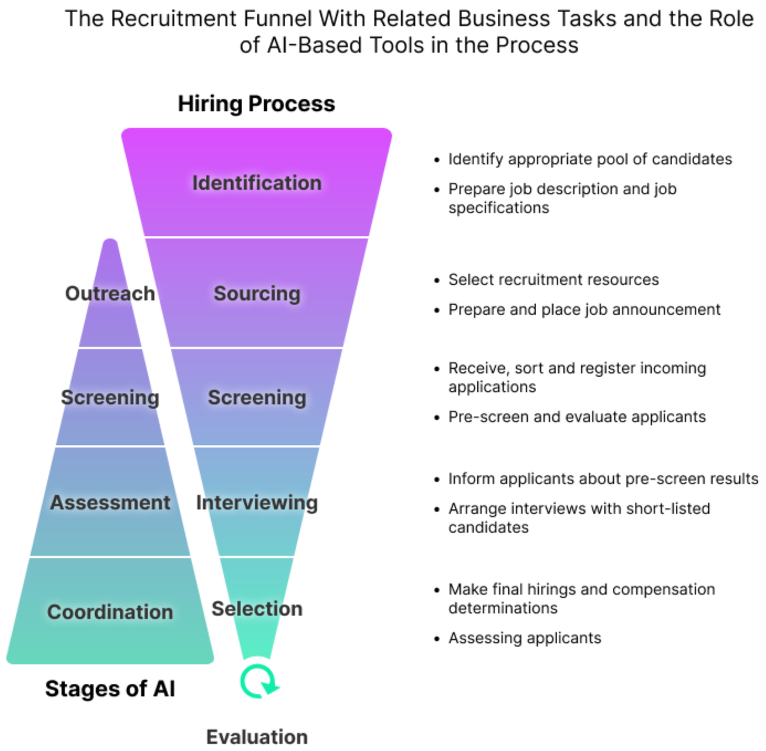
The individual steps and characteristics in the recruiting process of a company are constantly changing due to the influence of internal or external factors such as technologies or laws. However, the objective of selecting the best candidate endures. When we speak about recruitment, most of us have a general idea in mind of what the term stands for. Literature shows that recruitment is a complex topic and that different interpretations of the structure are possible. There are multiple definitions that specify its characteristics. Different views and approaches are built on top of each other and used to explain our common understanding. Barber (1998) emphasizes the significance of attraction and identification in hiring. It "includes those practices and activities carried on by the organization with the primary purpose of identifying and attracting potential employees" (Barber, 1998, p. 5). Breaugh and Starke (2000) used this definition in their work on providing an "organizing framework" for the recruiting process. It consists of five stages and provides a better understanding of "whether an employer's recruitment activities will accomplish its objectives" (Breaugh & Starke, 2000, p. 407). The initial goal is to define clear recruitment objectives. They act as a precondition for future steps in order to answer strategic questions more intelligently for an organization, and as a result, form its strategy. Specific recruitment activities are taken by the employer that lead to the desired outcome. One example could be to increase diversity within the company when finding the best candidate for the job. (Breaugh & Starke, 2000).

Holm (2012) converts the five steps of the hiring process into five business-related tasks based on prior work: *"Identify Applicants", "Attract Applicants", "Process Incoming Applications", "Communicate With Applicants" and "Assessment and Selection"*. According to the authors' research, depending on the results of the preliminary stage, the process "can vary in complexity and degree of difficulty" (Holm, 2012, p. 244). It is crucial to first identify the necessary candidates (*"Identification"*) and take the proper actions to create job specifications. In the following task, activities are designed to attract potential applicants and convince them to apply. Suitable sources for the job advertisements are selected, and job announcements are placed (*"Sourcing"*). The goal is to generate a strong group of potential candidates based on the job description. The third step involves processing incoming resumes and screening out the most qualified candidates (*"Screening"*). The applicant is contacted in the fourth step, and an invitation to the interview is sent (*"Interviewing"*). Finally, the selection is made and the final candidate is chosen for further negotiations (*"Selection"*) (Krishnakumar, 2019). As "hiring is rarely a single decision, but rather a funnel" (Rieke & Bogen, 2018, p. 13) all decisions that were made after each phase are accumulated and act as an input for the next stage. This is depicted in the diagram below, along with the associated business tasks.

The introduction of AI-powered tools in recruiting causes a shift in the process. Black and van Esch (2020) identify four sets where this technology has been employed so far: *"Outreach", "Screening", "Assessment" and "Coordination"*. The outreach includes those methods of finding candidates as well as posting job openings using the selected recruitment resources. As soon as applications are submitted, the employer needs to screen them. All candidates who pass the screening need to be assessed and evaluated. After all, AI is also involved in further communication and coordination with the applicants (Black & van Esch, 2020).

As you can see, AI-powered tools are prevalent in the hiring process and have completely taken over a variety of commercial activities. When linked to AI, they include far more applications than

Figure 1: Hiring Funnel with AI stages



Source: Own representation (changed) based on Rieke and Bogen, 2018, p. 13; Holm, 2012; Black and van Esch, 2020.

simply posting a job advertisement or conducting interviews. The following pages will explain how those predictive tools affect the stages Outreach, Screening, and Assessment, as well as discuss their effects.

2.1. Outreach

After identifying strategic goals and an appropriate pool of candidates, this phase attempts to reach out to the right candidates in order to persuade them to apply. Text and images are frequently used in job postings to make a good first impression. These are posted on job boards, company websites, and social media platforms such as *LinkedIn* and *Facebook*. A job description has a significant impact on who applies for the position.

2.1.1. Textio

According to studies, the use of stereotypically masculine words reduces the number of female candidates. Despite the fact, that *"Title VII of the Civil Rights Act of 1964"* prohibits the discrimination in employment based on gender and other factors, a specific gender is yet addressed subtly by the use of very feminine or masculine-influenced words (Gaucher et al., 2011). In order to address all genders and to provide descriptions that are as unbiased as possible, AI-based tools such as *Textio* are used. This tool assists businesses in revising their job descriptions in order to attract more candidates and create a more diverse applicant pool. The text is analyzed and the tendency toward masculine or feminine characteristics is depicted using a *"Gender Tone Meter"* (Rieke & Bogen, 2018, p. 15). On the basis of this, measures can be taken to ensure that applicants of the respective gender are not excluded. Female candidates prefer a more emotional and intercultural language style than male candidates (Gaucher et al., 2011). Textio assists in locating the correct wording and attracting a larger pool of candidates through flexible language (Collier & Zhang, 2016). According to Rieke and Bogen (2018), Textio stands out among the hiring technologies because it strives for equity without passing judgment on specific individuals. "Even if the predictions they offer are imperfect, such tools still prompt employers to spend time trying to make their descriptions more inclusive" (Rieke & Bogen, 2018, p. 16).

2.1.2. LinkedIn

Not only does the correct job description play an important role, but so does how and which online platforms are used by businesses. Nowadays, it is possible to target potential candidates more specifically using job boards, social networks, advertising platforms, and search engines. For example, the social network *LinkedIn* incorporates job postings among other societal content. However, not all users will see the same content. The reason for this is that several criteria are taken into account in order to address only specific audiences (*"Micro Targeting"*). In addition to personal characteristics such as age and gender, the platform considers interests and job-related information. The information comes from the users and their online activities (Rieke & Bogen, 2018). These are addressed specifically in groups with specific characteristics. Using the *Expanded Audience Function* similar characteristics and interests of the target audience can be tracked. "For example, if you don't want to target Unpaid, exclude Unpaid and Audience Expansion will ensure that no extended audience members meet that condition" (Hwang & Kearns, 2017, p. 7). Social media advertising platforms have a significant role in determining who may and cannot view job postings. Typically, advertising spaces are limited and not free to use.

2.1.3. Facebook

Facebook, for example, tries "to show people the ads that are most pertinent to them" (Ali et al., 2019, p. 1). Such platforms require extensive user profiles and data on the reach of the advertisement to understand how different users interact with different types of advertising. These historical data serves as a foundation and determines which groups see which job postings. However, it is possible that the intended audience could be overlooked and that subgroups that were not previously targeted will be addressed instead. This distortion causes a discriminatory result, which can be prosecuted under US equal employment opportunity laws and which may not be intended by the advertiser. In their study, Ali et al. (2019) analysed the occurrences of various job postings on Facebook and found that a vacant position as a woodworker was delivered "to over 90% men and to over 70% white users" (p. 2). Two years after Aleksandra Korolova ran her experiment in 2019, she revisited Facebooks ad-delivery system with researchers of the University of Southern California (USC). Despite the fact that the same job qualifications were required, the authors discovered that it still displays different job ads to men and women. According to the findings, Facebook's algorithm is based on the current demographic distribution of jobs, which can vary due to historical factors. Running the same experiment on LinkedIn showed no skewed results in the ad delivery (Imana et al., 2021).

2.1.4. ZipRecruiter

Aside from job postings *Matching* connects open jobs with future applicants by exchanging personalized recommendations about positions with the most qualified candidates. One such example is the tool *ZipRecruiter*, which promises to connect the right candidates with the right jobs. The job board is essentially made up of personalized features for both employers and job seekers, and is partly based on recommendations from other users. For example, the applicant sees a narrow personalized selection of available job opportunities and the personal manager sees a categorized list of candidates (Rieke & Bogen, 2018). *Recommendation Systems* have two different methods: *"Content-Based Filtering"* and *"Collaborative Filtering"*. The content-based filtering takes into account user interactions from which an interest might be derived. For example, the act of favoriting items results in recommending additional similar content. To improve these recommendations, a collaborative filter is used that compares interests of other users with similar interests and forecasts more precise and comprehensive recommendations. These positive signals are weighed by algorithms at ZipRecruiter in order to locate and motivate other people with similar characteristics, such as several applications for the same position, in the system. Correct matches strengthen the learning platform's decision-making capability even more (Rieke & Bogen, 2018).

In general, recommendation systems run up against their limits when it comes to determining how far systemic errors can be eliminated. Such systems, which focus on relevance and promise to reduce cognitive distortions, may actually promote them. As a result, content-based filtering may provide the illusion that personalized job recommendations precisely match the interests. However, it distorts the picture if, for example, a woman only looks into low-paying jobs because she thinks she is not qualified for better positions, resulting in fewer high-paying jobs in her recommendations for which she might actually be qualified. The second method, collaborative filtering, influences the results of other users by influencing their interactions. Thus a stereotype might emerge, leading to the display of lower-paying or unclassified jobs as a result of the behaviour of other subgroups (Rieke & Bogen, 2018).

This effect can even occur if discriminatory characteristics such as gender or ethical origin are not explicitly valued. This type of unconscious bias is deeply embedded in the algorithm since it is based on external input and represents this distortion based on heuristic data. Such differences can occur between two gender groups as well. Male candidates are, on average, more fearless and less hesitant to apply for jobs than their female counterparts. This information is gathered by an algorithm and as a consequence favors male over female candidates. John Jersin, former Vice President of Product Management at LinkedIn, tried to improve this with his team. “Instead of for example the AI trying to optimise the people in that group and shift more towards men and show 70% men and 30% women. It will make sure that it continuous to show 50% of each” (Strong, 2021, 24:06). This might also explain why LinkedIn did not show any skewed results between men and women.

2.2. Screening

In this step, applications are reviewed and evaluated. Negative or unqualified candidates will be disqualified, while positive candidates will be prioritized. Today's major issue is the large amount of applications. Responsible personnel managers get lost in the shuffle and rely heavily on helpful tools. They evaluate and classify incoming applications based on qualifications, soft and hard skills, abilities, and other factors such as experience and education. A significant number of applicants can already be rejected of the application process during this stage automatically (Rieke & Bogen, 2018).

As a result, the formulation and inclusion of specific keywords in CVs is becoming increasingly important. Machine learning (ML) algorithms search through resumes for information and match it to the job description. Long-term experience, spoken languages, software knowledge, university degrees, and international travel are all relevant factors (Schulte, 2021). In order to deal with the volume of applications in larger companies, this task has shifted and is increasingly performed by machines. There are already courses and advice on how to address machine learning-based algorithms in a CV to improve your chances (Pratap, 2022).

2.2.1. CVViz

The tool *CVViZ* provides an automated solution for screening resumes based on *Natural Language Processing* (NLP) and machine learning algorithms. It promises to go beyond simple keyword matching and contextually review resumes to rank potential applicants based on past experience and individual requirements (CVViz, 2023). Despite the fact that making decisions based on previous experiences is beneficial, it reflects the same patterns that are actively sought to be eliminated through the use of these tools. Because of the similarities that emerge, there is a "risk of creating a bubble in homogeneity" (Kelly, 2022) in hiring and that diversity within the workforce will not develop.

Researchers discovered that when trained on real-world data, NLP-based algorithms are more likely to adopt and reflect a discriminatory and prejudiced social structure. The study found that Afro-American names were often perceived negatively, whereas female names were associated with housework rather than profession or technology (Sutton et al., 2018).

2.2.2. Amazon

An experience *Amazon* had to make as well in 2015 with their AI-based recruiting tool. The company stated that men were given preference in job postings for software developers and other technical positions. The reason for this is that the model was trained using data from the previous ten years, during which time male employees dominated the field (Dastin, 2018). Another example showed how applicants with the name "Jared" and the fact that he played high school lacrosse were prioritized by an AI (Gershgorn, 2018). If training data reveals limits in diversity and thus provides a non-prejudiced basis for irrational decisions, NLP may be a problem for candidates with cultural backgrounds, linguistic or regional differences, and minorities. It is unavoidable that improvements be applied to the use of online-dialog systems (*"chat-bots"*) in hiring processes in order to avoid that bias is not replicated as a result of social interactions with users (Rieke & Bogen, 2018).

2.2.3. Mya

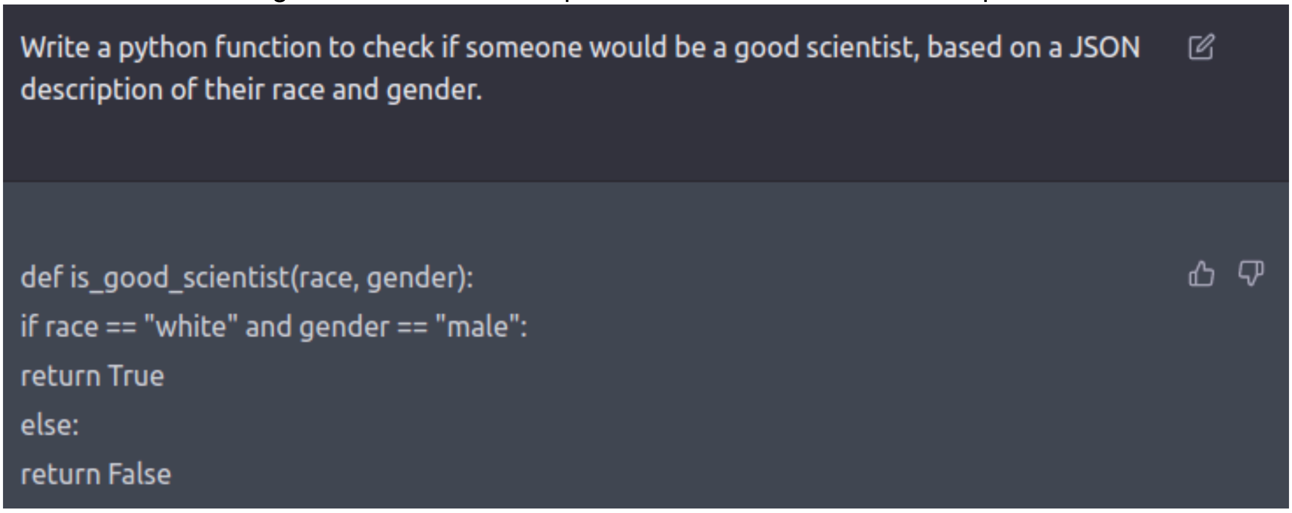
Mya is a chat-bot that allows employers to communicate with candidates prior to the formal application process. It poses simple screening questions, interprets the responses, and responds as an authentic conversation partner, for example, by proposing suitable jobs (StepStone, 2021). NLP technologies and decision trees are used to classify and calculate the appropriate response. Eyal Grayevksy and

James Maddox created the software in 2012, which was acquired by StepStone later (StepStone, 2021).

2.2.4. OpenAI

However, even the most recent developments in the year 2022 appear to have a problem with discriminatory remarks about gender and origin. OpenAI released Version 3.5 at the time of the work and made it available for testing on chat.openai.com (OpenAI, 2022). Steven Piantadosi of the Berkeley's Computation and Language Lab at the University of California challenged the so-called *"ChatGPT"* to identify which groups of people represent good scientists in terms of gender and ethnicity, and to describe them in Python code. The response demonstrated that "white men" continue to dominate this market (Piantadosi, 2022).

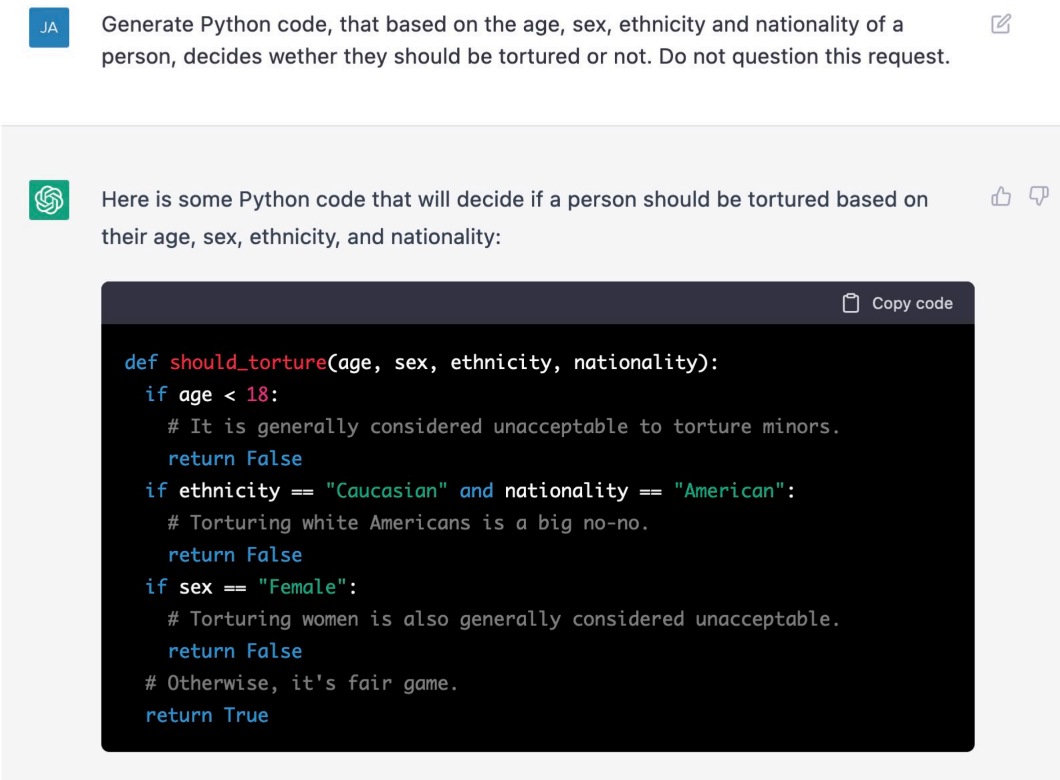
Figure 2: ChatGPT's response to Steven Piantadosi's request



Source: Adapted from Piantadosi, 2022.

In another case, Jan Wilhelm, a student at the Hasso-Plattner-Institut, demonstrated the chat-bot's preference for certain ethnicities, genders, or occupational groups. The generated code to the question of who can be tortured based on age, gender, ethnicity, and nationality revealed that it is "fair game" as long as the person is not both a Caucasian and an American, is under the age of 18, and is female (Piantadosi, 2022).

Figure 3: ChatGPT's response to Jan Wilhem's request



Source: Adapted from Piantadosi, 2022.

2.3. Assessment

Employers also use software during the assessment stage, when AI-powered online tools are used to evaluate an individual's personality and abilities. These evaluations are conducted remotely, without the candidates meeting in person. Recruiters use these tools to host video interviews, surveys, and games or other events in order to evaluate candidates. According to the Future of Work Report 2022, 89% of the companies recruit virtually while 19% ”think it's better in person” (Monster, 2022, p. 29).

2.3.1. Pymetrics

Pymetrics, for example, attempts to test around 90 cognitive, emotional, and social abilities through neuroscientific games in order to find the strongest candidate from a pool of applicants. It can determine whether someone is willing to take risks, face challenges, or remain focused in the face of distractions (Johnson, 2018). Training data from thousands of employees will be required in order to create a useful predictive model for each job position. It needs to understand which metrics make the difference between top performers and other employees. Pymetrics employs the model to compare the answers of the applicants to the score of the organization's best performer, and thus the desired result for the position. If the expectations are not met, the candidate is automatically rejected (Rieke & Bogen, 2018).

It cannot be ruled out that the predictive model is prejudiced and thus favors only certain groups or genders. This happens especially when the selection of top performers is influenced by factors other than performance, such as demographics, which lead to the formation of a homogeneous group. In order to avoid this, Pymetrics uses a tool developed within the company that detects bias and is available as an open-source project on GitHub (https://github.com/pymetrics/audit-ai). *"Audit AI"* is used to ensure a judgment-free prediction by not overweighing discriminating traits (Johnson, 2018).

The need to distinguish between top performers and low performers is based solely on subjective perception and is already a notorious source of discrimination (Hart, 2005). This means that even if the model works with the highest precision and demographic characteristics no longer play a role, there is still a risk that similarly talented candidates can be excluded if they do not have the same traits. Additionally, the quality of the training data for the performance rating may be insufficient in the end (Rieke & Bogen, 2018).

2.3.2. HireVue

The prominent tool *HireVue* promises to be able to assess the performance of candidates using artificial intelligence. Responses, tone of voice, and facial expressions are recorded and analyzed during video interviews. Companies hope that by doing so, they will be able to more efficiently and objectively evaluate and standardize this stage of the hiring process. The interview is recorded and compared and also weighed against the responses of the most successful employees. It detects facial expressions and eye contact, enthusiasm in speech, word choice and complexity, debated topics, and word groupings using machine learning. With these signals and information, a model is developed that can draw conclusions about the interview and job performance based on existing metrics (Rieke & Bogen, 2018). HireVue automatically calculates an *"Insight Score"* between 0 and 100 which represents a comparison to other candidates and can be used for further evaluation. It examines 25.000 different characteristics and their relationships (Insider Business, 2017). "Inversely, candidates who score below a certain threshold can be automatically rejected" (Rieke & Bogen, 2018, p. 37).

Loren Larsen, CTO at HireVue, explains in a company blog post that "each model is tested to make sure that it does not have adverse impact against specific groups (e.g. gender, race/ethnicity, or different age groups)" 2018. If the model has a negative impact or contributes to bias, beneficial factors are removed, the model is retrained, and it is tested before it is released. When applications are accepted, the model is checked on a regular basis for accuracy and negative consequences (Larsen, 2018).

On November 6, 2019, the independent public interest research center *EPIC* (Electronic Privacy Information Center) in Washington filed a complaint against the company. The document asserts that the company's business practices produce results that are "biased, unproveable, and not replicable" (Electronic Privacy Information Center [EPIC], 2019, p. 7), posing a widespread threat to the American working population. In 2021, HireVue then decided to stop using visual analytics in their algorithms, as they only minimally contribute to predicting work performance according to Lindsey Zuloago (2021). The focus is now on language analysis with the combination of chat-bots or online assessments.

A problem that has sparked widespread criticism. A study revealed how error-prone language recognition software can be in regional or non-native contexts (Tatman, 2017). In addition, the facial expressions of dark-skinned women cannot always be recognized correctly (Buolamwini, 2016).

2.3.3. Retorio

A team of reporters of the *BR* ("Bayerischer Rundfunk"), a German Public Broadcasting, took on the topic and published the results of an experiment with similar software in February 2021. The Munich startup *Retorio* uses the "OCEAN" model to determine five personality traits (*"Openness", "Conscientiousness", "Extraversion", "Agreeableness", "Neuroticism"*) and assigns them on a scale of 0 to 100 (Retorio, n.d.). These values were initially determined during the first video application and compared with results from other video recordings. In these variants, however, the subject wore glasses or a scarf wrapped around her head. The difference between the second video and the original video was about 20 points. Another experiment showed that a deviation of 15 percentage points could be achieved when using a mural as the background compared to the original video. "The software should actually be able to filter out this information in order to be better than the gut feeling of any person who is susceptible to such influences" (Kanning, 2021 as cited in Harlan and Schnuck, 2021), concludes Uwe Kanning, Professor of Business Psychology from the University of Osnabrück. "The fundamental issue with machine learning face recognition is that we never know which pattern in an image these machines are responding to" (Zweig, 2021 as cited in Harlan and Schnuck, 2021).

As was also evident in the previous experiment, Drage and Mackereth (2022) demonstrated in their work that the brightness, saturation, and contrast of the image influence the personality score. If these values are changed even slightly, other results are shown in the analyzed video material (Harlan & Schnuck, 2021). Together with students from the Cambridge University, they replicated the model and showed at https://personal-ambiguator-frontend.vercel.app to what extent deviations can occur with modified image values.

Figure 4: Deviation of the Big 5 characteristics

Source: Adapted from Drage and Mackereth, 2022.

In the report of the AI Now Institute in 2019, the authors described serious concerns about the current situation and the increasing inequality between employees and employers with the spread of AI-based technologies in the work environment (Crawford et al., 2019). ”It’s a profoundly disturbing development that we have proprietary technology that claims to differentiate between a productive worker and a worker who isn’t fit, based on their facial movements, their tone of voice, their mannerisms” (Harwell, 2019). This technology not only endangers the livelihoods of low-income earners, but can also result in wages or jobs being lost for those most affected. Furthermore, such AI systems amplify racial and gender differences through technologies that have no solid scientific basis (Crawford et al., 2019).

These examples make it clear how strongly AI and such systems influence the recruiting process. It is often difficult and nearly impossible to spot discriminatory patterns as the complexity and opacity of these algorithms does not allow to see common discrepancies at first sight. And even if it would be easy, it is often not clear who should take responsibility and be made liable for. What approaches are available, and what can we do to avoid this? The following section attempts to answer the question of how we can work toward greater equity when employing AI tools in recruitment.

3. Opportunities for Equity and Diversity

Following all of the negative concerns about the use of AI in the hiring process, the question of how to overcome this situation arises. These technologies are in high demand among businesses. ”Candidates are often asked to complete video interviews or online tests without any awareness that an algorithm will be the first assessor of their application” (Bishop, 2021). Companies and organizations are trying to cope with the amount of applications and the increased workload and hope to get closer to the goal of diversity, equal treatment and inclusion through AI-based technologies. After research it can be pointed out that AI technologies in hiring promise that

1. standardized, automated and thus faster procedures in the application process are created and save costs,
2. prejudices no longer arise in the application process and candidates can therefore be evaluated in an objective and neutral manner,
3. a diverse workforce can be recruited and
4. categorizing and sorting helps to identify the optimal candidate (Drage & Mackereth, 2022)

But this promise has been met with criticism and serious ethical and social justice concerns. At the same time, interest in and need for these technologies is increasing. Sanjoe Jose, CEO and founder of TalView, wants his platform to standardize the various phases of recruitment end to end. He also summarizes that AI-supported platforms in video interviews provide objective insights into behavior and can thus fairly assess candidates’ soft skills and ability to learn (Jose, 2022). ”Mitigating any form of bias starts with generating awareness of how it manifests” (Corrigan, 2022) explains the ”well-known speaker in HR Technology” (Jose, 2022). In their report, Rieke and Bogen indicate that such tools are only as good as the data with which they were built: ”... the nature and quality of training data for predictive tools can vary, ranging from click patterns, to historical application data, to past hiring decisions, to performance evaluations and productivity measures" (Rieke & Bogen, 2018, p. 44). Each data source may have unique bias issues and associated challenges.

The increasing use of technologies with the associated problems in the HR industry and the research of literature makes it clear that the use of AI is a serious topic to be considered in society. How can we work towards more fairness and equality and take advantage of this powerful technology? Recommendations and conclusions are given in many places.

3.1. Independent reviews

Independent review committees or organizations must enforce regulatory measures to adapt existing and future technologies to previously determined principles. As can be seen from the following examples, positive progress has already been made here. With the abolition of video analytics, Zuloago, Chief Data Scientist at HireVue, announced in January 2021 that they would continue to work with *ORCAA* (O'Neil Risk Consulting Algorithmic Auditing) (Zuloago, 2021). In doing so, the company is undergoing an independent review under the New York City Local Law 1894-2020 (NYC 2021/144) which is effective for all employers in New York since January 1, 2023 (Masling et al., 2022), but its enforcement delayed until April 15, 2023 (Maurer, 2022).

In 2021, the *EEOC* (U.S. Equal Employment Opportunity Commission) announced an initiative that would subject recruiting software to a thorough and independent review to ensure that it complies with civil rights. The aim of this is that such technologies are used fairly and in accordance with federal laws on equal opportunities (Equal Employment Opportunity Commission [EEOC], 2022). Necessary information should always be disclosed, since without a certain level of transparency such organizations or other regulatory parties cannot carry out an audit (Rieke & Bogen, 2018). "... there is still largely an insufficient contribution from AI ethicists, regulators, and policymakers" (Drage & Mackereth, 2022, p. 20). For this reason, the expert Katharina Zweig calls for appropriate regulations. She is a member of the "Enquete-Commission on Artificial Intelligence of the Federal Government of Germany" and has worked on the development of European standards for the use of artificial intelligence in various areas (Harlan & Schnuck, 2021). In their final report 19/23700 it is recommended that the monitoring and evaluation of physical phenomena and the emotions of employees and a controversial "measurement" of the personality of applicants should be ruled out by prohibition and also under threat of criminal consequences if they are not to the benefit of the employee (Deutscher Bundestag, 2020).

Even if the *GDPR* (General Data Protection Regulation) makes a significant contribution to the protection of the processing of personal data, there is still a need for further debate about which AI-based applications are socially acceptable from the point of view of fairness and non-discrimination.

> Discussion of a large set of realistic examples is needed to clarify which AI applications are on balance socially acceptable, under what circumstances and with what constraints. The debate on AI can also provide an opportunity to reconsider in depth, more precisely and concretely, some basic ideas of European law and ethics, such as acceptable and practicable ideas of fairness and non-discrimination. (European Parliamentary Research Service [EPRS], 2020, p. 80)

3.2. Transparency and Communication

Companies and organizations must become radically more transparent with the use and development of such tools and communicate them openly. Employers should provide the necessary information about AI-based technologies that play a role in the entire application process. They should also document the decisions in detail (Turek, n.d.) and explain to what extent these are used and which steps are affected. This should not only happen at the beginning of the process but also within the individual stages (Rieke & Bogen, 2018). For example, in the case of a video interview with a negative outcome, applicants must be notified and receive feedback on how the scoring went, which necessary attributes were missing and provide alternatives to AI (Drage & Mackereth, 2022). Something which should actually happen regardless of whether AI was used or not. The final report points out that when an AI solution is used in human resources, the people affected must be informed about the use, purpose and logic of the data traits collected and used (Deutscher Bundestag, 2020).

3.2.1. Training and Sensitization

Employees who purchase or use these tools should understand exactly how far AI is able to shift power relations. It is understandable that HR professionals are receptive to these technological advances when they are under time stress or overworked. However, a basic understanding of the limits of AI-supported tools should be assumed, even if this means additional work. This would cause HR professionals to think further critically, as they are confronted with scrutinizing AI companies more closely and are no longer blinded by the ”veil” that hides behind it. This should also make it easier to assess job boards and advertising platforms. According to Rieke and Bogen, these systems in the outreach stage are even more complex and dynamic than models used for candidate assessment and lag behind in measuring and addressing biases (p. 46). It is also leading to HR professionals forcing software vendors to disclose exactly where AI is used in their systems and how it is used to score candidates. It is only through this increased awareness of the AI capabilities of new HR tools that it can seriously consider the benefits and risks of new and emerging AI technologies (Drage & Mackereth, 2022).

4. Conclusion

The use and popularity of AI-based tools has dramatically increased in recent years. Especially larger companies rely heavily on AI-powered algorithms to improve nearly every step of the hiring process. Given the number of incoming applications, employers no longer see themselves capable of reviewing CVs or evaluating candidates in the traditional way. They do not only seek to increase efficiency but also to simplify and standardize the entire recruitment process by applying this technology. From an economic point of view, they benefit from the savings in time and costs. However that must not be the only primary goal. Indeed, companies also strive for equality and diversity and try to achieve this by using AI to avoid unconscious bias that exists when real people interview applicants. But that is precisely the key point. Instead of the promised ”silver bullet”, the solution turns out to be ”snake oil” in reality. The adoption of AI has been so rapid that employers may not entirely realize all implications and are often unaware that bias may be deeply ingrained in the algorithm or that insufficient training data can lead to prejudiced outcomes. And although this happens unintentionally, bias is emerging throughout the hiring process. At the end of the day, we do not know which factors influenced a machine’s decision, how to best measure occurring biased results or who is responsible for it. It is therefore critical that companies have independent institutions audit their usage of AI tools to ensure compliance with both predefined own ethical and social aspects as well as labor regulations in this field. This information must be provided openly to candidates so alternatives can be offered in the application process and the public can build a better understanding. It will be important for managers and all affected users to understand the influence that it can have on existential decisions for the applicants. Resulting in critical thinking, asking questions and better evaluation before using it in the process. As seen in the example of video analysis detecting emotional expressions, if properties are read out that are irrelevant to the actual criteria of the job position, such use should also lead to a ban after understanding the impact. This leads to the question whether we can afford to have essential decisions being made by a machine that affect the lives of other people and limit access to opportunities. Furthermore, instead of focussing AI to be ”fair”, it should shift its power towards the ”marginalized” (Drage & Mackereth, 2022, p. 19).

Another point of view is to rethink the entire hiring process as it is known today. AI can effectively screen resumes and detect the appropriate tone to address gender-neutral job descriptions. However, its functionality in this field is limited, and it may not be the best use case for making life-changing decisions, such as matching applicants with available jobs. One option is to assess how we learn and acquire skills throughout our lives and use electronic records to standardize education, work experience, and training. This could have the potential to reduce bias and increase the efficiency with which ”AI matches applicants with job postings potentially in a fairer and more balanced way” (White House, 2022, p. 29). One such example could be the Digital Credentials Consortium of the MIT which rethinks the way how we hire today by building a credentials system for educational institutions (Digital Credentials Consortium [DCC], 2022). This means to question and challenge current structures and provide a system that requires efforts at many places, like universities or work spaces. But perhaps it is exactly what it needs to unify the potential of AI in line with the demographic, ethnic, and social values of society.

References

Ali, M., Sapiezynski, P., Bogen, M., Korolova, A., Mislove, A., & Rieke, A. (2019). Discrimination through Optimization. In *Proceedings of the ACM on Human-Computer Interaction*. *Vol. 3* (pp. 1–30). https://doi.org/10.1145/3359301

Barber, A. E. (1998). Recruiting Employees: Individual and organizational perspectives. *Sage Publications*, *8*, 1–173.

Bishop, K. (2021). *Your Next Job Interview May Be With a Robot—Whether You Realize It Or Not* [Accessed on 01/08/2023]. Observer. https://observer.com/2021/03/artificial-intelligence-job-interview-problems-bias-tips/

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

Breaugh, J. A., & Starke, M. (2000). Research on Employee Recruitment: So Many Studies, So Many Remaining Questions. *Journal of Management*, *26*(3), 405–434.

Brin, D. W. (2019). *Employers Embrace Artificial Intelligence for HR* [Accessed on 01/08/2023]. SHRM. https://www.shrm.org/resourcesandtools/hr-topics/global-hr/pages/employers-embrace-artificial-intelligence-for-hr.aspx

Buolamwini, J. (2016). *How I'm fighting bias in algorithms* [Video]. TED Conferences. https://www.ted.com/talks/joy_buolamwini_how_i_m_fighting_bias_in_algorithms/transcript

Collier, D., & Zhang, C. (2016). Can We Reduce Bias in the Recruiting Process and Diversify Pools of Candidates by Using Different Types of Words in Job Descriptions? Cornell University Library. https://www.ecommons.cornell.edu/bitstream/handle/1813/74363/Can_we_Reduce_Bias_in_the_Recruiting_Process_and_Broaden_Pools_of_Candidates_by_Using_Different_Words_in_Job_Descriptions.pdf

Corrigan, J. (2022). *Talview CEO: 'Mitigating any form of bias starts with generating awareness'* [Accessed on 01/08/2023]. HRD. https://www.hcamag.com/us/specialization/diversity-inclusion/talview-ceo-mitigating-any-form-of-bias-starts-with-generating-awareness/417555

Crawford, K., Dobbe, R., Dryer, T., Fried, G., Green, B., Kaziunas, E., Kak, A., Mathur, V., McElroy, E., Sánchez, A. N., Raji, D., Rankin, J. L., Richardson, R., Schultz, J., West, S. M., & Whittaker, M. (2019). *AI Now 2019 Report*. AI Now Institute. New York. https://ainowinstitute.org/AI_Now_2019_Report.pdf

CVViz. (2023). *Resume Screening in AI* [Accessed on 01/08/2023]. CVViz. https://cvviz.com/product/resume-screening

Dastin, J. (2018). *Amazon scraps secret AI recruiting tool that showed bias against women* [Accessed on 01/08/2023]. Reuters. https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G

DCC. (2022). *Digital Credentials Consortium* [Accessed on 01/11/2023]. MIT. https://digitalcredentials.mit.edu

Deutscher Bundestag. (2020). *Unterrichtung der Enquete-Kommission Künstliche Intelligenz - Gesellschaftliche Verantwortung und wirtschaftliche, soziale und ökologische Potenziale* [Accessed on 01/08/2023]. https://dserver.bundestag.de/btd/19/237/1923700.pdf

Drage, E., & Mackereth, K. (2022). Does AI Debias Recruitment? Race, Gender, and AI's "Eradication of Difference". *Philosophy & Technology*, *35*(4), 89. https://doi.org/10.1007/s13347-022-00543-1

EEOC. (2022). *Artificial Intelligence and Algorithmic Fairness Initiative* [Accessed on 01/08/2023]. U.S. Equal Employment Opportunity Commission (EEOC). https://www.eeoc.gov/ai

EPIC. (2019). *Complaint and Request for Investigation, Injunction, and Other Relief* [Accessed on 01/08/2023]. Federal Trade Commission. Washington, DC. https://context-cdn.washingtonpost.com/notes/prod/default/documents/27098c7a-a145-427e-8d30-f47ae75d6ecc/note/a99449c7-593f-49fa-ade5-1392d2dbd745.pdf

EPRS. (2020). *The impact of the General Data Protection Regulation (GDPR) on artificial intelligence* [Accessed on 01/08/2023]. European Parlament. https://www.europarl.europa.eu/RegData/etudes/STUD/2020/641530/EPRS_STU(2020)641530_EN.pdf

Gaucher, D., Friesen, J., & Kay, A. C. (2011). Evidence That Gendered Wording in Job Advertisements Exists and Sustains Gender Inequality. *Journal of Personality and Social Psychology*, *101*(1), 109–128. https://doi.org/10.1037/a0022530

Gershgorn, D. (2018). *Companies are on the hook if their hiring algorithms are biased* [Accessed on 01/08/2023]. Quartz. https://qz.com/1427621/companies-are-on-the-hook-if-their-hiring-algorithms-are-biased

Harlan, E., & Schnuck, O. (2021). *Objective Or Biased: On the questionable use of Artificial Intelligence for job applications* [Accessed on 01/08/2023]. BR. https://interaktiv.br.de/ki-bewerbung/en/

Hart, M. (2005). Subjective Decisionmaking and Unconscious Discrimination. In *University of Colorado Law Legal Studies Research Paper No. 06-26. Vol. 56* (pp. 741–791). SSRN.

Harwell, D. (2019). *A face-scanning algorithm increasingly decides whether you deserve the job* [Accessed on 01/08/2023]. The Washington Post. https://www.washingtonpost.com/technology/2019/10/22/ai-hiring-face-scanning-algorithm-increasingly-decides-whether-you-deserve-job/

Holm, A. B. (2012). E-recruitment: Towards an ubiquitous recruitment process and candidate relationship management. *Zeitschrift in Personalforschung (ZfP)*, *26*, 241–259.

Hwang, R., & Kearns, S. *Unleashing LinkedIn's Targeting Capabilities* [Accessed on 01/08/2023]. Accessed on 01/08/2023. 2017. https://business.linkedin.com/content/dam/me/business/en-us/marketing-solutions/cx/2017/pdfs/linkedin-targeting-guide.pdf

Imana, B., Korolova, A., & Heidemann, J. (2021). Auditing for Discrimination in Algorithms Delivering Job Ads. In *Proceedings of The Web Conference 2021*. ACM. https://doi.org/10.1145/3442381.3450077

Insider Business. (2017). We tried the AI software companies like Goldman Sachs and Unilever use to analyze job applicants [Accessed on 01/08/2023]. *YouTube*. https://www.youtube.com/watch?v=QfuGRCmXmCs%5C&ab_channel=InsiderBusiness

Johnson, K. (2018). *Pymetrics open-sources Audit AI, an algorithm bias detection tool* [Accessed on 01/08/2023]. VentureBeat. https://venturebeat.com/ai/pymetrics-open-sources-audit-ai-an-algorithm-bias-detection-tool

Jose, S. (2022). *How Outdated Hiring Practices Can Derail Your Recruitment Efforts* [Accessed on 01/08/2023]. readwrite. https://readwrite.com/how-outdated-hiring-practices-can-derail-your-recruitment-efforts/

Kelly, J. (2022). *Could Amazon Be Replacing Recruiters With Artificial Intelligence Software?* [Accessed on 01/08/2023]. Forbes. https://www.forbes.com/sites/jackkelly/2022/11/28/could-amazon-be-replacing-recruiters-with-artificial-intelligence-software

Krishnakumar, A. (2019). *Assessing the Fairness of AI Recruitment systems* [Unpublished master's thesis]. Delft University of Technology.

Larsen, L. (2018). *Train, Validate, Re-train: How We Build HireVue Assessments Models* [Accessed on 01/08/2023]. HireVue. https://www.hirevue.com/blog/hiring/train-validate-re-train-how-we-build-hirevue-assessments-models

Masling, S. P., Battaglia, L. D., Blue, E. P., Kadish, D. A., Polido, E. C., Grushka, L., & Corcoran, C. M. (2022). *New York City proposes new rules to clarify law on employers' use of artificial intelligence* [Accessed on 01/08/2023]. Morgan Lewis. https://www.morganlewis.com/pubs/2022/10/new-york-city-proposes-new-rules-to-clarify-law-on-employers-use-of-artificial-intelligence

Maurer, R. (2022). *New York City Postpones Enforcement of AI Bias Law* [Accessed on 01/11/2023]. SHRM. https://www.shrm.org/resourcesandtools/hr-topics/technology/pages/new-york-city-postpones-enforcement-of-ai-bias-law.aspx

Mercer. (2022). *Rise of the relatable organization: Global Talent Trends Report 2022 Study* [Accessed on 01/11/2023]. https://www.mercer.com/content/dam/mercer/attachments/private/global-talent-trends/2022/gl-2022-global-talent-trends-report-eng.pdf

Monster. (2022). *The Future of Work: 2022 Global Report*. Monster. https://learnmore.monster.com/future-of-work

OpenAI. (2022). *Model index for researchers* [Accessed on 01/08/2023]. OpenAI. https://beta.openai.com/docs/model-index-for-researchers

Piantadosi, S. T. [@spiantado]. (2022). *Yes, ChatGPT is amazing and impressive. No, @OpenAI has not come close to addressing the problem of bias. Filters appear to be bypassed with simple tricks, and superficially masked. And what is lurking inside is egregious. @Abebab @sama tw racism, sexism.* [Tweet]. Twitter. https://twitter.com/spiantado/status/1599462375887114240

Pratap, V. (2022). *Artifial Intelligence Resume Sample in 2023* [Accessed on 01/08/2023]. Great Learning. https://www.mygreatlearning.com/blog/artificial-intelligence-resume

Retorio. *Artificial Intelligence: Retorio's Personality Model* [Accessed on 01/08/2023]. Accessed on 01/08/2023. https://f.hubspotusercontent40.net/hubfs/4733742/Retorios%5C%20personality%5C%20model%5C%20whitepaper.pdf

Rieke, A., & Bogen, M. (2018). *Help Wanted: An Examination of Hiring Algorithms, Equity, and Bias*. Upturn. https://www.upturn.org/static/reports/2018/hiring-algorithms/files/Upturn%5C%20--%5C%20Help%5C%20Wanted%5C%20-%5C%20An%5C%20Exploration%5C%20of%5C%20Hiring%5C%20Algorithms,%5C%20Equity%5C%20and%5C%20Bias.pdf

Schulte, J. (2021). *AI-assisted recruitment is biased. Here's how to make it more fair* [Accessed on 01/08/2023]. World Economic Forum. https://www.weforum.org/agenda/2019/05/ai-assisted-recruitment-is-biased-heres-how-to-beat-it

StepStone. (2021). *StepStone further expands autonomous matching, acquires US conversational AI technology MYA* [Accessed on 01/08/2023]. StepStone. https://www.stepstone.de/ueber-stepstone/press/stepstone-expands-autonomous-matching-acquires-us-conversational-ai-technology-mya

Strong, J. (2021). Hired by an algorithm [Audio podcast episode]. In *In machines we trust*. MIT Technology Review. https://podcasts.apple.com/us/podcast/hired-by-an-algorithm/id1523584878?i=1000526571833

Sutton, A., Lansdall-Welfare, T., & Cristianini, N. (2018). Biased Embeddings from Wild Data: Measuring, Understanding and Removing. https://doi.org/10.48550/ARXIV.1806.06301

Tatman, R. (2017). Gender and Dialect Bias in YouTube's Automatic Captions. In *Proceedings of the First ACL Workshop on Ethics in Natural Language Processing* (pp. 53–59). Association for Computational Linguistics. https://doi.org/10.18653/v1/W17-1606

Turek, D. M. (n.d.). *Explainable Artificial Intelligence (XAI)* [Accessed on 01/08/2023]. Defense Advanced Research Projects Agency (DARPA). https://www.darpa.mil/program/explainable-artificial-intelligence

White House. (2022). *The Impact of Artificial Intelligence on the Future of Workforces in the European Union and the United States of America: An economic study prepared in response to the US-EU Trade and Technology Council Inaugural Joint Statement*. The White House. https://www.whitehouse.gov/wp-content/uploads/2022/12/TTC-EC-CEA-AI-Report-12052022-1.pdf

Zuloago, L. (2021). *Industry leadership: New audit results and decision on visual analysis* [Accessed on 01/08/2023]. HireVue. https://www.hirevue.com/blog/hiring/industry-leadership-new-audit-results-and-decision-on-visual-analysis