How Does AI Contribute to Existing Bias in the Technological Hiring Process?

A Research Paper submitted to the Department of Engineering and Society

Presented to the Faculty of the School of Engineering and Applied Science
University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements for the Degree
Bachelor of Science, School of Engineering

Emily Buerk

Spring 2023

On my honor as a University Student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments

Advisor

Bryn E. Seabrook, Department of Engineering and Society

Introduction to AI and Hiring Bias

The computer science field is increasingly powerful in today's world, but it is marked by a lack of diversity that has been perpetuated by the very technology that has come out of the industry. Since the invention of more powerful computing tools and complex artificial intelligence algorithms there has been a push to integrate these tools into many different fields such as image processing, navigation, recommendation algorithms, and hiring tools. The tech industry severely lacks women and minority programmers and executives alike, even with growing efforts to reduce this hiring gap through the use of "unbiased algorithms." In order to understand the way that machine learning (ML) technology affects lack of diversity in the hiring process, it is imperative to acknowledge the sexism and prejudice that evolved with the tech industry. In the social construction of technology (SCOT) lens, humans created a diversity-lacking tech industry and used technology with biased data to unknowingly exacerbate their own discriminatory hiring practices. Artificial intelligence (AI) is complex, loosely monitored, and reliant on human data which contributes to the growing issue behind lack of diversity in the technological hiring process.

Research Question and Methods

In order to understand more about SCOT, gender bias in the technological hiring process, and how they can fit together to help society improve technology, this report primarily utilized news articles, government reports, and research papers on the subjects. News articles from Harvard Business Review, Bloomberg, and other credible sources are found by sifting through recent articles to discover the latest in "hiring bias," "artificial intelligence applications," "social construction of AI," "automated hiring tools," "natural language processing," "data collection and bias," and other similar keywords. ArXiv, a Cornell research website, and Google scholar are incorporated to find papers on “the social construction of AI,” “automated resume reviewing,” and “data bias effects on AI.” Together, these sources are organized both by category and chronologically by event type and timeline, respectively to tell the individual stories of these subjects and bring these themes together to understand AI hiring bias in the technological field with an STS lens. These resources aid in answering the question: How does AI contribute to the existing bias in the technological hiring process?

Social Construction of Technology and Hiring Bias

As the creators of technology, humans influence how tech is used and how it is made. Technology creators are not removed from the consequences of their inventions. In the Social Construction of Technology theory, “technology design is an open process” and the different outcomes it may produce are “depending on the social circumstances of development” (Klein & Kleinman, 2002). The second component to SCOT is relevant social groups. Relevant social groups is the idea that all members of a particular social group give the same meanings attached to specific artifacts (Klein & Kleinman, 2002). People have largely influenced the culture of the tech industry through the adoption of AI and the mutual understanding humans have shaped for the uses for AI.

SCOT also features a third component of closure and stabilization when an engineering problem is properly addressed by the technical solution designed for the job (Klein & Kleinman, 2002). In the case of AI, humans have not achieved a state of stabilization on this new technology since it continues to be shaped into new forms to address continuing issues in the world. The fourth SCOT component of wider context that explains the larger sociocultural and political components of the group who creates the technology (Klein & Kleinman, 2002). In the case of AI, the wider context symbolizes the historic cultural bias against the inclusion of underrepresented minorities in the technological industry.

When it comes to artificial intelligence, humans affect this technology due to the bias that seeps into the machines. When AI develops bias and humans choose to utilize these algorithms for important processes such as hiring tools, this perpetuates the effects of their creations (Matos et al., 2022). From the gender biased industry standard to the gender-targeted marketing for personal computers, people have constructed the gender-gap that plagues the tech industry. This gender bias has made its way into hiring technology and contributed to lower percentages of women in tech positions. Men hold roughly 80% of executive positions in the tech industry whereas women are only 20% (US EEOC, *Diversity in high tech*). This gender gap already leads to the shutting out of diverse perspectives, and lack of diversity is further exacerbated by the fact that minority groups are hired at much lower rates in tech as well (US EEOC, *Diversity in high tech*).

Lack of diversity in the tech industry is particularly negatively influential because of the increasing impact of tech on the world. Narrowing the voices of women and minorities in the field has shown a direct impact on the technological innovation that comes out of the field. Virtual reality headsets designed primarily by men were found to make many women feel nauseous (Chakravorti, 2021). Facial recognition algorithms created by an industry where 83.3% of executives are white have been known to discriminate against minority populations (US EEOC, *Diversity in high tech*). As described in SCOT theory, humans are the drivers of technology and its biases.

Cases of AI Bias

Major tech companies like Amazon and Google have been accused of using sexist AI in their hiring process. The Google lawsuit following these accusations cost $118 million and will not be stopping anytime soon (Thomas, 2022). Not only was Amazon using this technology inadvertently, but so was Google's job advertisement AI. Google's targeted advertisement system makes men much more likely to see high-paying executive jobs as opposed to women of equivalent qualifications (Simonite, 2020). Men were shown these jobs at a rate almost 6 times higher than women at equal level (Carpenter, 2015).

Google is the largest search engine in the industry and completes roughly 3.5 billion searches a day (*Google Search Statistics*). These Google searches have a large impact on the information that society will see. This can affect education, guide ideals, and influence career paths. While a single Google search may feel harmless, it can largely guide the trajectory of one's life. When women go to apply for jobs and see the applications for lower-paying parts while men automatically are directed to better positions, this would be very limiting. Based on trends in the tech industry with women making up an even smaller portion of the workforce than other careers, this gap is likely to be worse for women searching for tech positions.

Origins of Bias in Tech

The gender bias in the tech industry did not start with technology, though. People are behind the sexism that is being reflected by the biased artificial intelligence programs. The tech industry has been heavily biased since right after computers were invented in 1983-84. When the percentage of women in the industry went from 37% to 18%, it was not a coincidence (Fessenden, 2014). Hiring shifted to follow the System Development Corps sexist vision of an ideal computer scientist: the man who sits alone and codes in a dark room. Recent reports also show women feeling the negative effects of bias in the workplace which leads them to exit their roles early (US EEOC, *Diversity in high tech*). This male-centric vision of technology began fueling personal computers marketed to young boys rather than girls (Fessenden, 2014). This only led to further the cycle of men dominating the tech industry.

However, the tech industry bias is not limited to just gender. Underrepresented minority populations are severely lacking in the tech industry as well. Roughly 9% of the nation's top university computer science graduates are underrepresented minorities, yet they make up only 5% of the current workforce at large tech companies (US EEOC, *Diversity in high tech*). Had tech companies followed the pure talent, the issue of lack of diversity would be solved (Chakravorti, 2021). Research shows that diversity is a driver of market growth. Companies with diverse leadership are 45% more likely to have a growing market (Hewlet et al., 2014). On the contrary, companies with less diversity were found to squander innovative ideas from minority populations and lose out on the potential of their workforce (Hewlet et al., 2014). This issue of lack of diversity in corporations began with human bias, and at one point artificial intelligence was a promising way to remove biased human views from the equation. However, in recent years, the diversity gap has only been exacerbated by artificial intelligence. Essentially, computers do not actually learn to think “independently” or “objectively.” Machine learning algorithms all process data gathered by humans to “learn” to make complex decisions.

Engineered Inequity in Algorithm Training

Algorithm training is complex and involves multiple different stages of which all can contain bias. First, algorithms are trained with a goal in mind. In supervised, unsupervised, or semi-supervised learning the goal is either to make a decision or to analyze features of the data. The main issue that lies in this stage is how humans frame the goal of the algorithm (Hewlet et al., 2014). For example, if a recruiting manager asks the hiring AI to find candidates that are similar to their current employees, and their company is an overwhelmingly male-dominated tech force, the algorithm will mainly look for more men just like the 2018 Amazon algorithm. AI programs are trained to look for patterns to make these decisions by analyzing mass amounts of data to create neural networks and discover trends. Although the process mimics that of a human decision-maker, the machine still lacks true cognitive thinking that can actively combat biases. The issue of framing is purely caused by humans, but the problem is perpetuated by the AI that the users are training.

Even when the problem is framed correctly, like when the algorithm is looking for pure technical talent, bias can still be introduced at the data collection stage. All methods of deep learning are trained with datasets, and much of their computations take place in the “hidden layer” of the neural network (Sarker, 2021). Some algorithms are fed very controlled data from a variety of diverse sources. However, plenty of these datasets gather information from the mass data available on the web. The “recency bias” describes the issue with 90% of the world's data only coming from the last couple of years (Chatfield, 2022). In the tech industry, skills and qualifications are constantly training, so recruiters are using recent data to evaluate candidates. When this data is based on companies like Google’s tech sector that was as low as 17% women as recently as 2014, the AI program does not have gender balanced data (Sullivan, 2021). Data bias in AI also comes into play when the dataset consists only of existing groups. This phenomenon often leaves out less privileged groups of people who have not had the opportunities to succeed yet (Hargittai, 2018). Utilizing current strong employee resumes to train an AI model shows bias against people who do not have access to the same resources as those who were already hired and does a poor job at taking into account the unique situations of these individuals.

Preparing the data is another way that humans and AI create and perpetuate bias. Allowing AI machines to consider things like zip codes, backgrounds, education, and other attributes that correlate more with prejudice than talent leads to bias (Hao, 2020). Allowing AI machines to pick up on the trends that do not correlate directly with worker competence allows for the program to show favoritism towards certain groups. Discrimination emerges quickly in hiring when bias factors are allowed into algorithms for the sake of seemingly better results. The results of these programs are only as great as they are shown to be in testing, and hiring managers do not have a way to evaluate and understand the opportunity cost of the talent they are missing by relying heavily on these programs to narrow their applicant fields.

Validation training only endorses the engineered inequity left in the AI hiring algorithms. These studies attempt to evaluate the performance of AI models, usually by comparing it to existing datasets similar to those used for training. The issue here is that these datasets rarely reflect the entire population and often conceal certain minority groups (Raghavan & Barocas, 2022). Validation testing of the algorithms reports great results for the majority of the populations they are tested on, but conceals that the testing always has mistakes assessing minority groups which is a smaller percentage of the testing population. In 2020-2021, black workers made up 13% of the workforce yet only held 4% of tech positions (Carlton, 2022; Russonello, 2021). Similarly, hispanic workers are 17% of the total workforce but held a mere 8% of tech jobs in 2017-2019 (Nadeem, 2021). Since the percentage to test on is so low and validation testing is neglectful of these groups in the evaluation stage, this is another way that humans and machines contribute to the ignorant bias in creating hiring algorithms.

Hiring AI in Society

The dangerous part of the bias that is programmed into AI is the impact that this bias will have on future consumers. In 2021, 86% of companies reported that AI is becoming "mainstream technology" to their company (McKendrick, 2021). The majority of these companies stated that they accelerated the use of AI due to the COVID-19 pandemic during the increased reliance of technology at the time (McKendrick, 2021). Without proper care in implementation, overusing technology can have counterproductive effects due to the complete lack of humans in the decision making process (MIT Tech Review, 2022). This lack of human supervision is what is most detrimental to AI utilized in the hiring process to narrow down candidates without keeping the technology in check.

AI is primarily used in hiring with natural language processing to scan resumes and computer vision to process video interviews through companies like HireVue (Zuloaga et al., 2018). In order to advertise jobs, search engines like Google cater individual user listings. While these practices are shown to promote bias, it is also true that these algorithms still have the potential to reduce human bias. Algorithms audited by Northeastern University professionals were shown to be rid of bias (Vanderford, 2022). Companies like HireVue actively reevaluate their programs and eliminate prejudiced parameters, effectively trading higher validation rates for less biased algorithmic results (Zuloaga et al., 2018). In these cases, AI bias is correctable and has potential to be less biased than people (Moschella, 2022). Still, it is important to understand how to capture and understand this bias in machines in order to combat the larger issue.

Limitations

This paper reflects on the effects of AI on hiring bias in the technology industry specifically. The studies shown are limited to understanding the new high tech industry and recent bias that has been reflected by the integration of artificial intelligence into hiring practices. Artificial intelligence began to be utilized for hiring around 2015 and therefore there are few statistics that can be attributed to the engineered inequity in the tech hiring process specifically. Many of the statistics used show the larger trend of bias in the industry and the current results with and without using AI for resume reviews and video interview screenings.

In order to expand on this project, future research would include direct analysis of AI bias versus human bias. Additionally, further exploration of tactics for reducing AI bias and healthily incorporating AI into the hiring process would benefit this proposal. ChatGPT and other AI programs can be used as a supportive tool for recruiters to evaluate candidates without promoting bias (Wood, 2022). Using methods described by the data bias manager at HireVue reduces the amount of prejudiced features evaluated by AI to promote healthier relationships with hiring tools (Zuloaga et al., 2018). These methods for reducing bias can be further explored and analyzed in future work.

Conclusion

Artificial intelligence is a powerful tool, and its presence in the tech workplace has completely altered the hiring field. While the early stages of this AI hiring technology has proven to exacerbate bias in the tech field, it is also true that well-monitored AI can be trained to have less bias than humans. The tech field will greatly benefit from increased diversity and perspectives incorporated into future technologies to create inclusive products for all populations. Likewise, artificial intelligence must be properly researched before it becomes integrated into more areas. With the rise of technology like ChatGPT, it is important not to underestimate the power that AI has and the harm it will cause.

References

Carlton, G. (2022, November 23). *Why the tech diversity gap continues to persist: Bestcolleges*. BestColleges.com. Retrieved March 22, 2023, from https://www.bestcolleges.com/bootcamps/guides/tech-diversity-gap-persists/

Carpenter, J. (2015, July 7). *Google's algorithm shows prestigious job ads to men, but not to women*. The Independent. Retrieved March 22, 2023, from https://www.independent.co.uk/tech/google-s-algorithm-shows-prestigious-job-ads-to-men-but-not-to-women-10372166.html

Chakravorti, B. (2021, February 2). *To increase diversity, U.S. tech companies need to follow the talent*. Harvard Business Review. Retrieved March 22, 2023, from https://hbr.org/2020/12/to-increase-diversity-u-s-tech-companies-need-to-follow-the-talent

Chatfield, T. (2022, February 24). *The trouble with big data? it's called The 'recency bias'*. BBC Future. Retrieved March 22, 2023, from https://www.bbc.com/future/article/20160605-the-trouble-with-big-data-its-called-the-recency-bias

Fessenden, M. (2014, October 22). *What happened to all the women in computer science?* Smithsonian.com. Retrieved March 22, 2023, from https://www.smithsonianmag.com/smart-news/what-happened-all-women-computer-science-1-180953111

Hao, K. (2020, April 2). *This is how AI bias really happens-and why it's so hard to fix*. MIT Technology Review. Retrieved March 22, 2023, from https://www.technologyreview.com/2019/02/04/137602/this-is-how-ai-bias-really-happens and-why-its-so-hard-to-fix/

Hargittai, E. (2018, July 30). *Potential biases in big data: Omitted voices on Social Media*. Social Science Computer Review. Retrieved March 23, 2023, from https://journals.sagepub.com/doi/10.1177/0894439318788322

Hewlet, S. A., Marshall, M., & Sherbin, L. (2014, August 1). *How diversity can drive innovation*. Harvard Business Review. Retrieved March 22, 2023, from https://hbr.org/2013/12/how-diversity-can-drive-innovation

Hewlet, S. A., Marshall, M., & Sherbin, L. (2014, August 1). *How diversity can drive innovation*. Harvard Business Review. Retrieved March 22, 2023, from https://hbr.org/2013/12/how-diversity-can-drive-innovation

Insights, M. I. T. T. R. (2022, August 9). *Embracing the rapid pace of ai*. MIT Technology Review. Retrieved March 22, 2023, from https://www.technologyreview.com/2021/05/19/1025016/embracing-the-rapid-pace-of-ai/

Internet Live Stats. (n.d.). *Google search statistics*. Google Search Statistics - Internet Live Stats. Retrieved March 22, 2023, from https://www.internetlivestats.com/google-search-statistics/

Klein, H. K., & Kleinman, D. L. (2002). *The social construction of Technology: Structural Considerations*. The Social Construction of Technology: Structural Considerations. Retrieved April 21, 2023.

Matos, E. J., Flores, T. F., & Bertoncini, A. L. C. (2022, May 31). *Technological determinism and the Social Construction of Technology*. Admethics. Retrieved March 22, 2023, from https://www.admethics.com/technological-determinism-and-the-social-construction-of-technology/

McKendrick, J. (2021, September 27). *AI adoption skyrocketed over the last 18 months*. Harvard Business Review. Retrieved March 22, 2023, from https://hbr.org/2021/09/ai-adoption-skyrocketed-over-the-last-18-months

Moschella, D. (2022). *Ai bias is correctable. human Bias? not so much*. RSS. Retrieved March 22, 2023, from https://itif.org/publications/2022/04/25/ai-bias-correctable-human-bias-not-so-much/

Nadeem, R. (2021, April 1). *Stem jobs see uneven progress in increasing gender, racial and ethnic diversity*. Pew Research Center Science & Society. Retrieved March 22, 2023, from https://www.pewresearch.org/science/2021/04/01/stem-jobs-see-uneven-progress-in-increasing-gender-racial-and-ethnic-diversity/

Raghavan, M., & Barocas, S. (2022, March 9). *Challenges for mitigating bias in algorithmic hiring*. Brookings. Retrieved March 22, 2023, from https://www.brookings.edu/research/challenges-for-mitigating-bias-in-algorithmic-hiring/

Russonello, G. (2021, June 21). *How big tech allows the racial wealth gap to persist*. The New York Times. Retrieved March 22, 2023, from https://www.nytimes.com/2021/06/21/us/politics/big-tech-racial-wealth-gap.html

Sarker, I. H. (2021, August 18). *Deep learning: A comprehensive overview on techniques, taxonomy, applications and research directions - SN computer science*. SpringerLink. Retrieved March 22, 2023, from https://link.springer.com/article/10.1007/s42979-021-00815-1

Simonite, T. (2020, April 2). *Probing the dark side of Google's ad-targeting system*. MIT Technology Review. Retrieved March 22, 2023, from https://www.technologyreview.com/2015/07/06/110198/probing-the-dark-side-of-googles-ad-targeting-system/

Sullivan, G. (2021, October 26). *Google statistics show Silicon Valley has a diversity problem*. The Washington Post. Retrieved March 22, 2023, from https://www.washingtonpost.com/news/morning-mix/wp/2014/05/29/most-google-employees-are-white-men-where-are-allthewomen/

Thomas, P. (2022, June 13). *Google settles gender discrimination lawsuit for $118 million*. The Wall Street Journal. Retrieved March 22, 2023, from https://www.wsj.com/articles/google-settles-gender-discrimination-lawsuit-for-118-million-11655079036

US EEOC. (n.d.). *Diversity in high tech*. Diversity in High Tech. Retrieved March 22, 2023, from https://www.eeoc.gov/special-report/diversity-high-tech

Vanderford, R. (2022, February 2). *Ai could cut hiring biases as companies make push to find workers, proponents say*. The Wall Street Journal. Retrieved March 22, 2023, from https://www.wsj.com/articles/ai-could-cut-hiring-biases-as-companies-make-push-to-find-workers-proponents-say-11643797804

Wood, M. (2022, November 23). *Ai used for hiring and recruitment can be biased. but that's changing.* Marketplace Tech. Retrieved March 22, 2023, from https://open.spotify.com/episode/1ov5lE4Troub67CAlLC4to?si=B-NdI_eBRJK0rgp15PYfLQ&nd=1

Zuloaga, L., Benson, C., & Whitenack, D. (2018, October 22). *Practical AI: Machine learning, data science: Fighting bias in AI (and in hiring) on Apple podcasts*. Apple Podcasts. Retrieved March 22, 2023, from https://podcasts.apple.com/us/podcast/fighting-bias-in-ai-and-in-hiring/id1406537385?i=1000422345962