Procedural Justice and Fairness in Automated Resume Parsers for Tech Hiring: Insights from Candidate Perspectives

Swanand Vaishampayan, Sahar Farzanehpour, Chris Brown
Department of Computer Science
Virginia Tech, Blacksburg, VA, USA
{swanandsv, saharfarza, dcbrown}@vt.edu

2023 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC) | 979-8-3503-2946-9/23/$31.00 ©2023 IEEE | DOI: 10.1109/VL-HCC57772.2023.00019

***Abstract*—To streamline tech hiring processes, the use of talent management platforms has emerged as a new norm. AI-driven Automated Resume Parsers (ARPs) are aimed to simplify the application process for candidates and employers. However, ARP designs typically prioritize employers over candidates. Further, prior work also demonstrates these AI systems are not able to achieve the intended goals of inclusivity and fairness for candidates, negatively impacting minorities in the tech hiring pipeline. Thus, aspiring IT professionals on the job market often spend significant time and effort preparing applications, only to have their resume rejected by an AI model without receiving attention from a recruiter. This work aims to study candidates' perspectives of ARPs. We sent an survey, receiving responses from 103 undergraduate and graduate CS students, and analyze their perspectives through a prism of *procedural justice*, a measure of fairness. By introducing procedural justice and opting for a human-centered design approach, we believe AI models in the hiring pipeline can achieve the intended goals of inclusivity and fairness. The findings from this study will be beneficial for future designs of more transparent and fair ARPs.**

***Index Terms*—tech hiring, automated resume parsers, procedural justice, bias in tech**

I. Introduction

The tech hiring pipeline comprises of—but is not restricted to—job ads, preliminary resume screening, an initial HR interview, round(s) of technical interviews, and job offers [9]. To streamline initial screening process and handle tasks related to human resources, companies often use talent management solutions, such as Workday [4] or Internet Collaborative Information Management Systems (ICIMS) [2]. These platforms typically include *automated resume parsers* (ARPs). ARPs are AI-powered natural language processing (NLP) models which automate the resume screening process. Features such as Automated Tracking System (ATS), where resumes are ranked based on their content, are often integrated into ARPs. According to JobScan [20], over 98% of the Fortune-500 companies use ATS in their hiring pipeline. A screenshot of an example ARP in the Workday application can be found in Appendix C Fig 2.

The purpose of ARPs is twofold: First, applicants on the tech job market typically send resumes to a wide range of open positions simultaneously [17]. ARPs can auto-fill the job application based on the resume of candidates, saving invaluable time. For employers, ARPs give an option for recruiters to auto-reject candidates based on application questions. Another key feature is 'Application Ranking' which typically uses 'keyword matching' which allows the company recruiter to save time by filtering the top ranked applications.

AI-driven hiring is also intended to improve equality and fairness in hiring processes by reducing human bias. However, prior work states otherwise [40]. For example, biases against women [29] and dismal increases of minority populations in the tech workforce [21] raise serious doubts over the equality imparted by ARPs. It is plausible that ARPs are biased, contributing to a "leaky pipeline" in tech hiring where qualified candidates-–often from underrepresented backgrounds-–fail to attain positions in the software industry [28].

One of the most prominent measures of fairness is *procedural justice*, or the fairness of processes used by those in authority to reach a certain decision [11]. ARPs can make hiring processes faster, but the perception of improving equality is questionable as AI models can be immensely biased [32]. Therefore, it is important to examine usage of ARPs from both ethical and procedural justice points of view. From an ethical standpoint, ARP usage is questionable as biased models can discriminate against minority candidates while ranking resumes. In terms of procedural justice, the use of AI is under scrutiny as it fails to support the four founding principles: equality, fairness, trust and transparency [19], [23]. Even though there are two primary stakeholders of hiring pipeline, *candidates* and *employers*, the latter are given preference while designing hiring platforms.

The candidate perspective on ARPs is largely unexplored. Therefore, this work seeks to understand candidates' perspectives of ARPs and how candidates feel about AI-based resume parsers. The research questions addressed during the study are:

RQ1: What are candidates' perspectives on ARPs with regards to procedural justice?

RQ2: How do candidates perceive ARPs?

RQ3: What features should an *ideal* automated resume parser for candidates demonstrate?

Answers to our RQs are sought through the analysis of survey responses from candidates applying to IT positions. The results from this study will help motivate a human-centered ARP design, where emphasis would be on human-in-the-loop systems, and provide a deeper understanding for designing a

1943-6106/23/$31.00 ©2023 IEEE
DOI 10.1109/VL-HCC57772.2023.00019

reliable, secure and trustworthy AI-based hiring system [35] with the goal of promoting equity, diversity, and inclusion in the tech workforce.

II. Related Work

A. ARP Design and Ethics in AI-driven Recruitment

Prior work has suggested designing ARP systems to rank candidates' resumes using various NLP techniques, including lexicon analysis [6], transformer models such as BERT [12], and named entity recognition (NER) [30]. A major challenge in ML systems is their unpredictability for generalization–if trained on bias sample, the model inherits biases and generates biased predictions. To tackle this, Yeargar et al. suggested routine bias audits [40]. Yet, Sloane et al. [37] show that even bias check audits of multi-stage models cannot guarantee an unbiased process. Deshpande et al. [14] developed a fair tf-idf mechanism to mitigate socio-linguistic barrier in ARP design.

Ethical concerns have been raised over AI-driven hiring. Zhou et al. [41] linked mistrust with the "black-box" nature of NLP algorithms. Previous studies also show modern AI-driven hiring fail to incorporate algorithmic equity for disabled candidates [39] and introduce concerns for data privacy, transparency, and accountability [22]. Fernández-Martínez and colleagues [15] studied the legal and ethical implications of AI-based interview analysis and ranking systems used by hiring platforms, such as HireVue [1] and ModernHire [3]. To overcome ethical issues, this work seeks to incorporate procedural justice in AI-driven hiring for ARPs.

B. Employer perspectives in AI hiring

Li et al. [27] found that AI tools save time and effort for recruiters–but this comes at an immense cost of lack of precise control, less data accuracy, and a mismatch in algorithmic results and recruiters' expectations. Laurim et al. [26] noted the concerns raised by recruiters over AI conducted analysis of video interviews. A study by Robinson [33] found that HR participants did not support the use of AI for interviewing, as AI will fail to match human empathy. Thus, previous work suggests AI models are not able to replicate human judgment. We aim to investigate AI-based ARPs from the perspective of job applicants to understand their perceptions and motivate the need for better and more trustworthy systems.

C. Bias in the Tech Hiring Pipeline

Prior work has found biases across different stages of the hiring pipeline. Böhm et al. [13] found a dominance of masculine adjectives and nouns in job postings by IT companies in Germany. Behroozi et al. [9] suggest the data structure-focused nature of technical interviews results in a bias for candidates with more time and resources to prepare [8]. Hall and Gosha [18] show that Computer Science students at HBCUs face anxiety during preparation for technical interviews. Further, current technical interview practices exert added pressure and stress on candidates [7], which results in a decline in performance that disproportionately affects female applicants [10]. Work by Shuy et al. [36] discussed first-impression biases. Our work focuses on the resume screening stage of the hiring process, and we aim to understand candidates' perceptions of AI-based systems and believe incorporating procedural justice into ARPs might be a first step in addressing these complex issues.

III. Methodology

A. Survey and Participants

A survey based study was designed to gain insight into tech job applicants' opinions and experiences with resume screening systems. An IRB was approved for the study. Basic information on ARPs was provided to participants before completing the survey. The survey contained demographic questions along with a combination of short-answer, open-ended, and Likert scale questions. Specifically, we were interested in investigating participants' views on incorporating procedural justice in AI-based ARPs. To that end, we asked questions focusing on the four pillars of procedural justice i.e. equality, fairness, trust and transparency. More information regarding participants can be found in Appendix B.

B. Data Analysis

Likert scale and multiple choice questions were analyzed using statistical tests. Open-ended questions were analyzed using an open-coding approach on a subset of responses ($n = 10$) for each free response question. Then, two independent researchers reviewed and coded a larger subset of data ($n = 64$) individually to determine the categories of responses. We calculated our inter-rater agreement using Cohen's Kappa regarding perceptions ($k = 0.69$) and experiences ($k = 0.76$) for ARPs in addition to their ability to reduce bias in hiring ($k = 0.71$) for "substantial" agreement [25]. Disagreements were resolved after discussions between the coders. The consensus of codes were applied to the remaining responses. We were particularly interested in how minority candidates view procedural justice in ARPs. We define minorities as non-male identifying participants for gender, and non-Caucasian and non-Asian participants for race/ethnicity [34].

IV. Results

TABLE I
Candidates' Perspectives on Procedural Justice in ARPs

Attribute	**Population**	***n* (%)**	**Statistic**	***p-value***
Equality and Fairness	Overall	55 (69%)	-	-
	Gender	29 (57%)	$0.3673(\chi^2)$	0.5445
	Race	18 (33%)	$0.1788(\chi^2)$	0.6724
Trust	Overall	24 (23%)	-	-
	Gender	11 (46%)	$0.1241(\chi^2)$	0.7247
	Race	5 (21%)	$2.411(\chi^2)$	0.1205
Transparency	Overall	37 (36%)	-	-
	Gender*	9 (25%)	31662 (t)	***0.001***
	Race*	2 (5.5%)	6.36 (t)	***< 0.00001***

*** denotes statistically significant results (**p-value < 0.05**)

A. RQ1: Procedural Justice

We analyzed candidates' viewpoints on the procedural justice of ARPs by asking questions on the equality, fairness, trust, and transparency of these AI-based systems. To further analyze procedural justice, we examined the impact of race and gender on perceptions of fairness in resume screening systems. Table I presents an overview of these results.

1) Equality and Fairness: Survey contained an open-ended question regarding bias to analyze candidates' perceptions of the equality and fairness of ARPs. Out Of the 80 valid responses, the majority of respondents answered that ARPs do reduce bias ($n = 55, 68.75\%$), and there was no statistically significant difference for participants from underrepresented genders ($\chi^2 = 0.3673$, $p = 0.5445$) or races ($\chi^2 = 0.1788$, $p = 0.6724$). Major themes discovered were removal of personal biases by anonymizing the process and reduction in human errors to provide fair chances to all.

However, a few responses ($n = 25, 31.25\%$) reported that ARPs do not remove the biases in tech hiring. The major areas of concern reported by participants were programmatic errors and bias introduced later. We provide guidelines to reduce programmatic errors in AI models for ARPs in Section V. Additionally, participants feared that the multi-stage nature of hiring can re-introduce bias later in the hiring pipeline. This finding is seconded by work conducted by Sloane et al. [37]. Example participant responses are provided in Appendix A1. This motivates future work to investigate ways to incorporate procedural justice in other stages of the tech hiring pipeline.

2) Trust: We asked participants about their *trust* in ARPs. Our results suggest candidates do not trust ARPs to screen applications, with 77% of participants preferring a human reviewer over an AI-based system(see Figure 1). We further examined these results to investigate whether minorities have less trust in AI-based ARP systems using the X^2 test. For gender, we found minority participants had a higher average of ARP distrust (82.5%) compared to participants identifying as male (78%). However, it was not statistically significant ($X^2 = 0.1241, p = 0.7247$). Meanwhile for race/ethnicity, only five underrepresented participants preferred ARPs and there was a large difference in the percentage of minority participants preferring humans over AI (86%) compared to those representing the majority (72%). This difference was not statistically significant ($X^2 = 2.411, p = 0.1205$).

3) Transparency: Participants answered a Likert-scale question aimed to understand perceptions of transparency. We found that most participants ($n = 37, 36\%$) agreed or strongly agreed with the statement that ARPs make the hiring process more transparent. This contradicts prior work, which suggests that AI-models lack transparency [19], [26]. However, upon closer examination, an unpaired t-test revealed that there was a significant difference in responses from minority participants based on race/ethnicity (t = 6.636, p < 0.00001) and gender ($t = 31662$, $p = 0.001$). This indicates that minority participants find automated resume parsing systems lack transparency–signifying a major difference in the perception of the transparency of ARPs based on candidates' background.

B. RQ2: Perception

The overall perception about current ARPs was found to be negative ($n = 53$). Our qualitative analysis gave us three major themes: Filtering, Keyword matching, and Parsing. The majority of candidates ($n = 43$) stated that ARPs were not able to parse their resume information correctly (see examples in Appendix A2). Other areas of improvements include: "language understanding" ($n = 5$), where ARPs are able to understand context, not just match keywords; a "human-in-the-loop approach" ($n = 6$), where rather than having fully automated systems, humans are in charge of the final decision; and "feedback" ($n = 15$) about the workings of ARPs, where candidates felt the black-box nature of ARPs leads to ambiguity. The only benefit of reported by participants ($n = 16$) was that ARPs save time and make application completion easier (see Appendix A2).

C. RQ3: Ideal Features

Participants answered a ranking based question to rank the features of ARPs based on their importance and an open-ended question to seek more useful features for an ideal ARP system.

1) Ranking: The most important feature from candidates' perspectives was anonymizing resumes, thereby only picking up prior experience and skills ($n = 39$). This was followed by the ability to parse different templates of resumes ($n = 22$) and more accurate parsing ($n = 13$).

2) Open-ended: The responses yielded some interesting findings. Useful features to improve ARPs include:

1) *Selective parsing*, where the candidate is able to select the parts which should be parsed by an ARP;
2) *Reporting the confidence of parsing* to increase the reliability of the ARP; and
3) *Automated profile evaluation* where ARPs rank applications considering other relevant profiles (i.e. LeetCode, GitHub, etc.).

V. Discussion

Our findings suggest minorities do not find ARP predictions transparent (Section IV-A3). Responses revealed instances where minority participants "*faced a lot of discriminatory issues*" (P74) in tech hiring. Previous work also demonstrates ML models can be biased based on gender [24], [29] and race/ethnicity [38]. Further, our findings show that the majority of candidates do *not* trust ARP systems, preferring a human over resume parsers (Section IV-A2), and the overall perception of ARPs is negative (Section IV-A1). Therefore, the tech hiring pipeline has a lot of room for improvement. Based on the findings of our study, we suggest guidelines to incorporate procedural justice in future designs of ARPs

A. Avoid Programmer Errors

Programmer errors represent aspects of ARP design which are in control of the programmer. These include training data and evaluation, which have also been underlined in prior work [29], [32], [40]. Failing to resolve these issues results in the generation of biased models, which can detrimentally effect the tech hiring pipeline.

1) Training Data: It is important to make sure training data is balanced and diverse to avoid training models to become biased against minorities. There are previous instances when imbalanced data has lead to biased model generation against woman [29]. Training on balanced datasets where every category has equal representation can minimize the training bias. In extreme cases of imbalanced dataset, sampling techniques such as oversampling and undersampling [31] and double bias correction [5] can help minimize the bias problem.

2) Evaluation: Evaluation of models is important in making sure systems do not make biased predictions. In terms of ARPs, this means evaluating the system on all resumes including those of minorities. [32] listed factors that could lead to biased model generation. To make the evaluation more robust, beta testing could be used to check the system against edge cases. In general, AI driven systems are hard to evaluate as it is almost impossible to replicate all the different edge cases when models are deployed in real time. Therefore, regular audits are important to incrementally improve the system. Another strategy that could be useful is evaluation of the system by minority participants.

B. Incorporate Explainability

Explainability is hard to achieve due to the sparse nature of language models. For ARPs, we can provide explainability by integrating example template prompts and feedback to users.

1) Example Prompts: As an ARP designer, a key focus should be on incorporating the ability to parse as many resume templates as possible. However, this is most likely infeasible due to the large space of potential resume formats and designs. Thus, an example prompt could help. For example, listing all possible templates which can be accurately parsed by the parser will help candidates mold their resumes according to the ideal templates for their job applications.

2) Feedback: Current ARPs do not contain information on back-end workings of the system. This makes ARP designs ambiguous for candidates. A major issue preventing this feedback is company policies. Nevertheless, we recommend to having at least an overview of the inner workings of the system to help users understand how their resumes are interpreted and analyzed by AI models.

C. Make Resistant to Data Perturbations

The major issue reported by candidates for current ARPs is that they work on keyword matching. Keyword matching can have a negative impact on resumes which do not contain the exact matching keyword, but rather a different word to express the same attribute. For example, JavaScript can also be written in short form as JS, “coding experience” could also be written as “programming experience”, etc. Another disadvantage of exact keyword matching is missing out on candidates with experience in similar technology stack. Therefore, context understanding is important for ARP designs. Modern transformer NLP models, such as BERT [12] and GPT-3 [16], would be well suited for future ARP designs.

D. Human-Centered Design

Our study findings show that the current fully automated design of ARPs might not be a good approach to employ in the tech hiring pipeline, as candidates distrust AI-based systems. Participants further stressed the need of a human reviewer for their resumes, rather than fully automated ARPs. Based on feedback from candidates, we believe a human-centered design approach for ARPs would be more suitable for the resume screening stage. High automation and high human control is the preferred quadrant for reliable, safe and trustworthy systems in the human-centered approach proposed by Shneiderman [35]. This would be achievable by incorporating features, like automatically anonymizing resumes, into ARPs. By only presenting the most relevant information from resumes to a recruiter (e.g. relevant skills and experience) with a human-in-the-loop approach, we can save time for HR workers and at the same time minimize personal biases and increase transparency in the resume screening process.

VI. Limitations and Future Work

While our participant sample was diverse, the results of our study may not generalize to all candidates seeking positions in the software industry. Our study was constrained to CS students, as they represent candidates on the job market pursuing tech careers and internships. However, student perception may not generalize to all candidates, such as experienced software engineers in job transition. Another limitation of this work is that Transgender is an independent identity, but was given as a separate option within gender demographics. Future studies can investigate different populations of candidates to understand their perceptions of ARPs.

There is a dearth of detailed information about ARPs used in the software industry. Most systems are proprietary, only accessible to company employees. Therefore, we believe these systems should be open sourced to promote anti-bias research on ARPs. Future work can explore the perspectives of employers to gain insight into the challenges faced by HR representatives using AI-based resume screening systems and to understand the feasibility of a human-in-the-loop approach. Another area for future research would be designing a more fair and transparent ARP. The guidelines from this study would be a good starting point for future designs.

VII. Conclusion

The use of AI models in resume screening processes is one area of concern in the tech hiring pipeline. Automated resume parsers, AI-based models used by employers to parse and rank resumes, can streamline reviews but also introduce bias. We investigate candidates’ perspectives of ARPs by surveying CS students on their perceptions of resume parsing systems in terms of procedural justice. We found participants distrust AI-based ARPs, and prefer human reviewers. Further, there was a significant difference in perceptions of the transparency by respondents from minority backgrounds. Based on our findings, we propose guidelines for future ARP designs to promote equity, diversity, and inclusion in the tech workforce.

References

[1] Hirevue. https://www.hirevue.com/.
[2] Internet collaborative information management systems (ICIMS). https://www.icims.com/.
[3] Modern hire. https://modernhire.com/.
[4] Workday. https://www.workday.com/.
[5] Sharmin Afrose, Wenjia Song, Charles B Nemeroff, Chang Lu, and Danfeng Yao. Subpopulation-specific machine learning prognosis for underrepresented patients with double prioritized bias correction. *Communications medicine*, 2(1):111, 2022.
[6] Cernian Alexandra, Sgarciu Valentin, Martin Bogdan, and Anghel Magdalena. Leveraging lexicon-based semantic analysis to automate the recruitment process. In *Transactions on Engineering Technologies: 25th World Congress on Engineering (WCE 2017)*, pages 189–201. Springer, 2019.
[7] Mahnaz Behroozi, Alison Lui, Ian Moore, Denae Ford, and Chris Parnin. Dazed: measuring the cognitive load of solving technical interview problems at the whiteboard. In *Proceedings of the 40th International Conference on Software Engineering: New Ideas and Emerging Results*, pages 93–96, 2018.
[8] Mahnaz Behroozi, Chris Parnin, and Titus Barik. Hiring is broken: What do developers say about technical interviews? In *Visual Languages & Human-Centric Computing (VL/HCC)*, pages 1–9, 2019.
[9] Mahnaz Behroozi, Shivani Shirolkar, Titus Barik, and Chris Parnin. Debugging hiring: What went right and what went wrong in the technical interview process. In *Proceedings of the ACM/IEEE 42nd International Conference on Software Engineering: Software Engineering in Society*, pages 71–80, 2020.
[10] Mahnaz Behroozi, Shivani Shirolkar, Titus Barik, and Chris Parnin. Does stress impact technical interview performance? In *Proceedings of the 28th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering*, ESEC/FSE 2020, page 481–492, New York, NY, USA, 2020. Association for Computing Machinery.
[11] Sarah Bennett, Lorelei Hine, and Lorraine Mazerolle. Procedural justice, 2018. https://www.oxfordbibliographies.com/view/document/obo-9780195396607/obo-9780195396607-0241.xml.
[12] Vedant Bhatia, Prateek Rawat, Ajit Kumar, and Rajiv Ratn Shah. End-to-end resume parsing and finding candidates for a job description using bert. *arXiv preprint arXiv:1910.03089*, 2019.
[13] Stephan Böhm, Olena Linnyk, Jens Kohl, Tim Weber, Ingolf Teetz, Katarzyna Bandurka, and Martin Kersting. Analysing gender bias in it job postings: A pre-study based on samples from the german job market. In *Proceedings of the 2020 on computers and people research conference*, pages 72–80, 2020.
[14] Ketki V Deshpande, Shimei Pan, and James R Foulds. Mitigating demographic bias in ai-based resume filtering. In *Adjunct publication of the 28th ACM conference on user modeling, adaptation and personalization*, pages 268–275, 2020.
[15] Carmen Fernández-Martínez and Alberto Fernández. Ai and recruiting software: Ethical and legal implications. *Paladyn, Journal of Behavioral Robotics*, 11(1):199–216, 2020.
[16] Luciano Floridi and Massimo Chiriatti. Gpt-3: Its nature, scope, limits, and consequences. *Minds and Machines*, 30:681–694, 2020.
[17] Manolis Galenianos and Philipp Kircher. Directed search with multiple job applications. *Journal of economic theory*, 144(2):445–471, 2009.
[18] Phillip Hall Jr and Kinnis Gosha. The effects of anxiety and preparation on performance in technical interviews for hbcu computer science majors. In *Proceedings of the 2018 ACM SIGMIS Conference on Computers and People Research*, pages 64–69, 2018.
[19] Michael Hind, Stephanie Houde, Jacquelyn Martino, Aleksandra Mojsilovic, David Piorkowski, John Richards, and Kush R Varshney. Experiences with improving the transparency of ai models and services. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–8, 2020.
[20] James Hu. Over 99% of fortune 500 companies use applicant tracking systems (ats). *Jobscan Blog*, 2019. https://www.jobscan.co/blog/99-percent-fortune-500-ats/.
[21] E Huet. Facebook's hiring process hinders its effort to create a diverse workforce. *Bloomberg Technology, January*, 9, 2017.
[22] Anna Lena Hunkenschroer and Christoph Luetge. Ethics of ai-enabled recruiting and selection: A review and research agenda. *Journal of Business Ethics*, 178(4):977–1007, 2022.
[23] Luyuan Jiang, Xin Qin, Xiaowei Dong, Chen Chen, and Wanqi Liao. How and when ai-human order influences procedural justice in a multistage decision-making process. In *Academy of Management Proceedings*, volume 2022, page 10419. Academy of Management Briarcliff Manor, NY 10510, 2022.
[24] Anja Lambrecht and Catherine Tucker. Algorithmic bias? an empirical study of apparent gender-based discrimination in the display of stem career ads. *Management science*, 65(7):2966–2981, 2019.
[25] JR Landis and GG Koch. The measurement of observer agreement for categorical data. *Biometrics*, 33(1):159—174, March 1977.
[26] Vanessa Laurim, Selin Arpaci, Barbara Prommegger, and Helmut Krcmar. Computer, whom should i hire?–acceptance criteria for artificial intelligence in the recruitment process. 2021.
[27] Lan Li, Tina Lassiter, Joohee Oh, and Min Kyung Lee. Algorithmic hiring in practice: Recruiter and hr professional's perspectives on ai use in hiring. In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society*, pages 166–176, 2021.
[28] Alexis Martin. *The leaky tech pipeline: A comprehensive framework for understanding and addressing the lack of diversity across the tech ecosystem*. Kapor Center for Social Impact, 2018.
[29] David Meyer. Amazon reportedly killed an ai recruitment system because it couldn't stop the tool from discriminating against women. *Fortune, October*, 10:2018, 2018.
[30] Vrinda Mittal, Priyanshu Mehta, Devanjali Relan, and Goldie Gabrani. Methodology for resume parsing and job domain prediction. *Journal of Statistics and Management Systems*, 23(7):1265–1274, 2020.
[31] Roweida Mohammed, Jumanah Rawashdeh, and Malak Abdullah. Machine learning with oversampling and undersampling techniques: overview study and experimental results. In *2020 11th international conference on information and communication systems (ICICS)*, pages 243–248. IEEE, 2020.
[32] Dena F Mujtaba and Nihar R Mahapatra. Ethical considerations in ai-based recruitment. In *2019 IEEE International Symposium on Technology and Society (ISTAS)*, pages 1–7. IEEE, 2019.
[33] Marcia F Robinson. *Artificial Intelligence in Hiring: Understanding Attitudes and Perspectives of HR Practitioners*. Wilmington University (Delaware), 2019.
[34] Allison Scott, F Kapor Klein, Frieda McAlear, Alexis Martin, and Sonia Koshy. The leaky tech pipeline: A comprehensive framework for understanding and addressing the lack of diversity across the technology ecosystem. Technical report, Technical report, https://www. kaporcenter.org/the-leaky-techpipeline-a …, 2018.
[35] Ben Shneiderman. Human-centered artificial intelligence: Reliable, safe & trustworthy. *International Journal of Human–Computer Interaction*, 36(6):495–504, 2020.
[36] Roger W Shuy. In-person versus telephone interviewing. *Inside interviewing: New lenses, new concerns*, pages 175–193, 2003.
[37] Mona Sloane, Emanuel Moss, and Rumman Chowdhury. A silicon valley love triangle: Hiring algorithms, pseudo-science, and the quest for auditability. *Patterns*, 3(2):100425, 2022.
[38] Latanya Sweeney. Discrimination in online ad delivery. *Communications of the ACM*, 56(5):44–54, 2013.
[39] Nicholas Tilmes. Disability, fairness, and algorithmic bias in ai recruitment. *Ethics and Information Technology*, 24(2):21, 2022.
[40] Lynette Yarger, Fay Cobb Payton, and Bikalpa Neupane. Algorithmic equity in the hiring of underrepresented it job candidates. *Online information review*, 44(2):383–395, 2020.
[41] Jianlong Zhou, Fang Chen, and Andreas Holzinger. Towards explainability for ai fairness. In *xxAI-Beyond Explainable AI: International Workshop, Held in Conjunction with ICML 2020, July 18, 2020, Vienna, Austria, Revised and Extended Papers*, pages 375–386. Springer, 2022.

APPENDIX

A. Example Participant Responses

1) RQ1: Procedural Justice: Table II below provides sample representative quotes from survey participants on the positives and negatives of ARPs with regard to equality and fairness. Fig. 1 shows the results regarding "Trust". We found that candidates trust human screening over ARPs.

TABLE II
POSITIVES AND NEGATIVES

Positives	
Anonymizing	**Fair Chances**
• *"Yeah, ARP can help to minimize unintended biases while reviewing a Resume or cover letter. For example, we can disable fields such as age, gender, school or university name, a candidate's photograph, and date of birth." (P2)* • *"I think a lot of bias can come from names, race, age, sex, gender, etc. and I think that an automated resume parser could remove these fields before passing them off for human review." (P8)*	• *"Reduce the errors and missing out of information by humans." (P15)* ‘ • *"They can reduce bias, by fairly giving chances to every qualified applicant." (P34)*

Negatives	
Programmatic Errors	**Bias in Later Stages**
• *"I think automated resume parsers are useful but ultimately, the ranking produced by the parser is based on the company's own biases and reflects qualities that the company values/rejects." (P1)* • *"I think that resume parsers will be exactly as Human biased as they are programmed to be." (P18)*	• *"They probably do but much of it will just be reintroduced later in the process" (P5)* • *"I don't think it will reduce bias, the recruiters are already unfamiliar with most of the candidates applying for the position and would not favor a single candidate. Moreover, if they really wanted to favor someone, the ATS would not stop them from accepting or rejecting the application." (P45)*

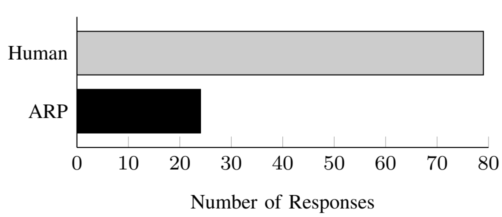

Fig. 1. Participants Responses on Trust for Resume Screening

2) RQ2: Perception: To analyze candidates' perception of ARPs, our survey consisted of an open-ended question asking what they like or dislike about them. The overall perception of ARPs was found to be negative.Table III presents example comments from participants on the most common responses.

TABLE III
PERCEPTION

Parsing Ability	**Time**
"No matter the format of resume, the parsers mess up the information on the resume and place the information in irrelevant places." (P9)	*"I like that they save me time when filling out applications and I like that they help me know how to communicate my skills more effectively on my resume."* (P6)
"It did not correctly work with a 2-column format." (P36)	*"I like that it lets me fill out an application faster since it reads through the resume."* (P13)

B. Participant Demographics

Table IV presents the detailed information regarding the demographics and, the industry experience of the participants of our study.

TABLE IV
PARTICIPANT DEMOGRAPHICS AND EXPERIENCE

Demographics		
	Percentage (%)	**Count (n)**
Gender		
Female	38%	39
Male	57%	59
Not specified	3%	3
Transgender	2%	2
Race		
Asian	54%	56
Hispanic or Latino	15%	15
African American	14%	14
Caucasian	12%	12
Multiracial	1%	1
Not specified	5%	5
Industry Experience		
Experience Level		
2-5 years	31%	
Less than 2 years	29%	
Internship only	19%	
No experience	13%	
Over 5 years	9%	
Automated Resume Parsers		
Familiar with concept	83%	85
Experienced in usage	72%	74

C. Resume Parser Example

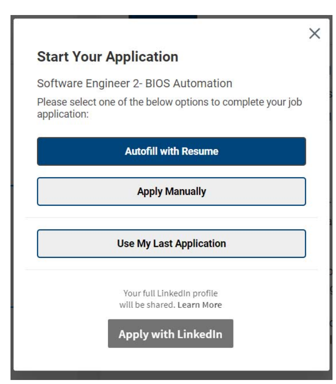

Fig. 2. Example ARP in Workday to autofill a resume