Social exclusion in personnel selection – The risk of discriminating AI biases

Clarissa Sabrina Arlinghaus
clarissa_sabrina.arlinghaus@uni-bielefeld.de
Center for Cognitive Interaction Technology
Work and Organizational Psychology
Bielefeld University
Bielefeld, North Rhine-Westphalia, Germany

Günter W. Maier
ao-psychologie@uni-bielefeld.de
Work and Organizational Psychology
Research Institute for Cognition and Robotics
Bielefeld University
Bielefeld, North Rhine-Westphalia, Germany

ABSTRACT

Work plays a central role in the life of adults as it opens up access to a wide range of valuable resources (e.g., financial security, time structure, social contacts). Thereby work contributes to the social inclusion of people in most societies. Therefore, personnel selection processes carry a high level of social responsibility. Nowadays, artificial intelligence (AI) is widely used in human resources (HR), but the unreflected use of AI in recruitment can lead to the exclusion of vulnerable groups. AIs are often trained with biased data which unconsciously results in discriminatory hiring practices. The aim of this workshop contribution is threefold. First, we distinguish between the generic term "social exclusion" and its subtype "discrimination". Second, we raise awareness for discrimination in AI-based personnel selection. Third, we aim to foster interdisciplinary discussions about the responsible use of AI-based selection processes to prevent the severe consequences of social exclusion (e.g., helplessness, depression, suicide).

CCS CONCEPTS

• **Applied computing** → **Psychology**; • **Computing methodologies** → **Intelligent agents**; • **Social and professional topics** → **Computer supported cooperative work**.

KEYWORDS

social exclusion, discrimination, AI, personnel selection

ACM Reference Format:
Clarissa Sabrina Arlinghaus and Günter W. Maier. 2023. Social exclusion in personnel selection – The risk of discriminating AI biases. In *Proceedings of Interdisciplinary Approaches in Human-Agent Interaction (Inter.HAI '23).* ACM, New York, NY, USA, 3 pages.

1 INTRODUCTION

Recent advancements in AI transform nowadays working environment. For example, AI is used for workforce scheduling [8], support in administrative HR tasks [1] and easier onboarding processes [24]. AI is also used to regulate access to scarce and valuable resources such as jobs [15, 19] or sensitive information within a company [25].

SAIL is an interdisciplinary and interinstitutional project in Germany that focuses on the full life-cycle of AI to ensure a sustainable long-term development. Within SAIL we explore processes of social inclusion and exclusion [33]. Social exclusion is defined as a complex and multidimensional process where people lack resources, rights, goods, services, or participation opportunities which affects not only their individual quality of life but also the society as a whole [23]. Social exclusion entails severe consequences like alienation, helplessness, depression [35] or even suicide [2, 27]. There are different types of social exclusion (e.g., silent treatment, dehumanization, discrimination), but they do not differ in their negative effects, as all excluded people feel threatened in their basic needs (i.e., belonging, self-esteem, control, meaningful existence) [34].

In this workshop contribution, we will focus on discrimination, which is one type of social exclusion [34]. Discrimination occurs when individuals are rejected or unfairly treated because of personal attributes such as age, gender, national origin, race, sexual orientation, disability, or any other factor and therefore lose options and opportunities [21].

AI-based personnel selection can increase recruitment quality and efficiency but can also lead to discriminating hiring practices [10, 22]. Biases in managerial decisions have disadvantaged minorities for many years [11]. Now, AI replicates these biases because it is trained with biased human decisions [31]. AI biases have been found in HR recruitment, selection, and development [22]. Such digital discrimination is a growing problem as more and more decisions are handed over to AI [13]. It also affects a range of disciplines (e.g., computer science, law, sociology) but has not been solved by any of them yet [13].

With this workshop contribution, we draw attention to the danger of social exclusion through discriminatory AI biases. We explain how social exclusion affects individuals of subjugated groups and how this is replicated through AI. This way, we contribute to a better understanding of the concepts "social exclusion" and "discrimination" as well as their consequences, also with respect to the rise of AI. For illustration, we use the example of AI-supported personnel selection. However, the transfer of personnel selection procedures to other selection processes is desirable and encouraged. In this workshop, we will discuss the problem of discriminating AI biases in selection processes as well as potential prevention strategies in an interdisciplinary manner acknowledging its broad relevance.

2 RELATED WORK

2.1 Social Exclusion

Over the years, numerous definitions of "social exclusion" have been presented, focusing on different aspects [6]. Although there

Inter.HAI '23, Dec 04, 2023, Gothenburg, Sweden

is no single accepted definition of social exclusion [26], researchers do agree that social exclusion has severe negative effects. For example, social exclusion promotes health inequalities, depressive symptoms [4], negative emotions, low self-esteem [14], low trust to others [29], bad psychological well-being [5, 17], infectious diseases, cardiovascular problems [3], and sleep disorders [16].

In the past, social exclusion was often only considered in interpersonal relationships. However, social exclusion through technical agents should not be neglected. We claim that the effects of social exclusions are as harmful in human-technology interaction as in human-human interaction. This assumption is supported by studies that observed that exclusion from computer players leads to a comparable worsening of belonging, self-esteem, control, and meaningful existence as exclusion from human players [20, 36]. Consequently, social exclusion by AI agents must be taken seriously.

A lack of participation is a core aspect of most definitions for social exclusion [26]. The unreflected use of AI can prevent disadvantaged groups from participating in society, although this is usually not the intention of the programmers [7]. We would like to illustrate this with the use case of AI-based personnel selection, where people could be prevented from participating in companies because of AI biases.

Unemployment is directly connected to multiple dimensions of social exclusion and causes detrimental effects on people's social status, economic resources, life satisfaction, self-efficacy, and mental health [32]. Reemployment reduces distress through gains in e.g., financial situation, status, time structure, collective purpose, and social contact [37]. Consequently, personnel selection processes play a crucial role in social inclusion.

2.2 AI in Personnel Selection

Today, companies increasingly rely on AI to make personnel selection faster, easier, and more efficient [15, 19]. In most cases, the software is only used for pre-selection (e.g., ranking candidates' CVs for further decision-making), which is then checked by humans working in HR [12]. However, this pre-selection can be biased [10, 22].

A bias is the inclination of an unfair decision for or against one person or group [28]. This creates discrimination [10, 22].

AI agents are not free from bias because AI was likely trained with biased data, biased measurement, biased human decisions, preference for majorities, non-representative samples, and missing values [30]. Nevertheless, many people believe in unbiased AI. As a result, people may prefer their job application to be evaluated by an algorithm rather than a human [31] or are blind to the biased behavior of a technical agent [18]. Discrimination through AI evokes less moral outrage because people do not assume a prejudicial motivation [9]. Instead, people attribute the discrimination experience to their own performance [18].

AI systems are often seen as "black boxes" where it is almost impossible to notice discrimination based on gender, race, nationality, or any other minority [15]. Thus, there is a high risk that certain individuals will be filtered out of AI-based selection processes at an early stage which affects individuals and society.

3 CLOSING REMARKS

Therefore, we address AI biases in personnel selection processes to shed light on this mostly unconscious but equally harmful type of social exclusion. This way, we hope to raise some awareness to this issue to encourage a more reflected usage of AI agents in selection processes.

To prevent the severe effects of social exclusion we recommend an interdisciplinary development of AI selection processes, as already suggested by [15]. We are interested in exchanging ideas and experiences with practitioners in human resources and AI development as well as researchers from various disciplines (e.g., computer science, sociology, business administration).

We look forward to an interdisciplinary discussion of our future research about discriminating AI biases in personnel selection processes. We are planning a series of online experiments to explore the effects of social exclusion in AI-supported personnel selection. Vignettes and questionnaires will probably be used. These are typical psychological methods, but we also welcome methodological suggestions from other disciplines.

ACKNOWLEDGMENTS

SAIL is funded by the Ministry of Culture and Science of the State of North Rhine-Westphalia under the grant no NW21-059A.

REFERENCES

[1] Marwan Mohamed Abdeldayem and Saeed Hameed Aldulaimi. 2020. Trends And Opportunities Of Artificial Intelligence In Human Resource Management: Aspirations For Public Sector In Bahrain. *International Journal of Scientific & Technology* 9, 1 (Jan. 2020), 3867–3871.

[2] Bushra Akram, Juwairya Nawaz, Zeeshan Rafi, and Abrar Akram. 2018. Social exclusion, mental health and suicidal ideation among adults with hearing loss: protective and risk factors. *JPMA. The Journal of the Pakistan Medical Association* 68, 3 (March 2018), 388–393.

[3] Robert W. Aldridge, Alistair Story, Stephen W. Hwang, Merete Nordentoft, Serena A. Luchenski, Greg Hartwell, Emily J. Tweed, Dan Lewer, Srinivasa Vittal Katikireddi, and Andrew C. Hayward. 2018. Morbidity and mortality in homeless individuals, prisoners, sex workers, and individuals with substance use disorders in high-income countries: a systematic review and meta-analysis. *Lancet (London, England)* 391, 10117 (Jan. 2018), 241–250. https://doi.org/10.1016/S0140-6736(17)31869-X

[4] Ana Maria Arias-Uriona and Natalie Guillén. 2020. Inequalities in depressive symptoms between natives and immigrants in Europe: the mediating role of social exclusion. *Cadernos de Sáude Pública* 36, 11 (Oct. 2020), e00124319. https://doi.org/10.1590/0102-311x00124319

[5] Gökmen Arslan. 2018. Social Exclusion, Social Support and Psychological Well-being at School: A Study of Mediation and Moderation Effect. *Child Indicators Research* 11, 3 (March 2018), 897–918. https://doi.org/10.1007/s12187-017-9451-1

[6] Carsten Kronborg Bak. 2018. Definitions and Measurement of Social Exclusion —A Conceptual and Methodological Review. *Advances in Applied Sociology* 08, 05 (May 2018), 422–443. https://doi.org/10.4236/aasoci.2018.85025

[7] Solon Barocas and Andrew D. Selbst. 2016. Big Data's Disparate Impact. *California Law Review* 104, 671 (Sept. 2016), 671–732. https://doi.org/10.2139/ssrn.2477899

[8] Dominik Bentler, Stefan Gabriel, Benedikt A. Latos, and Günter W. Maier. 2023. Prozessmodell personalorientierter Entscheidungsgestaltung einer künstlichen Intelligenz. *Personal Quarterly* 2 (March 2023), 18–23.

[9] Yochanan E. Bigman, Desman Wilson, Mads N. Arnestad, Adam Waytz, and Kurt Gray. 2023. Algorithmic discrimination causes less moral outrage than human discrimination. *Journal of Experimental Psychology: General* 152, 1 (Jan. 2023), 4–27. https://doi.org/10.1037/xge0001250

[10] Zhisheng Chen. 2023. Ethics and discrimination in artificial intelligence-enabled recruitment practices. *Humanities and Social Sciences Communications* 10, 1 (Sept. 2023), 567. https://doi.org/10.1057/s41599-023-02079-x

[11] George B. Cunningham and Harper R. Cunningham. 2022. Bias among managers: Its prevalence across a decade and comparison across occupations. *Frontiers in psychology* 13 (Nov. 2022), 1034712. https://doi.org/10.3389/fpsyg.2022.1034712

[12] Papiya Das, Manjusha Pandey, and Siddharth Swarup Rautaray. 2018. A CV Parser Model using Entity Extraction Process and Big Data Tools. *International*

Journal of Information Technology and Computer Science 10, 9 (Sept. 2018), 21–31. https://doi.org/10.5815/ijitcs.2018.09.03

[13] Xavier Ferrer, Tom van Nuenen, Jose M. Such, Mark Coté, and Natalia Criado. 2021. Bias and Discrimination in AI: A Cross-Disciplinary Perspective. *IEEE Technology and Society Magazine* 40, 2 (June 2021), 72–80. https://doi.org/10.1109/MTS.2021.3056293

[14] Jonathan Gerber and Ladd Wheeler. 2009. On Being Rejected: A Meta-Analysis of Experimental Research on Rejection. *Perspectives on Psychological Science* 4, 5 (May 2009), 468–488. https://doi.org/10.1111/j.1745-6924.2009.01158.x

[15] Manuel Gonzalez, John Capman, Frederick Oswald, Evan Theys, and David Tomczak. 2019. "Where's the I-O?" Artificial Intelligence and Machine Learning in Talent Management Systems. *Personnel Assessment and Decisions* 5, 3 (Nov. 2019), 33–44. https://doi.org/10.25035/pad.2019.03.005

[16] André Hajek and Hans-Helmut König. 2022. Loneliness, perceived isolation and sleep quality in later life. Longitudinal evidence from a population-based German study. *Archives of Gerontology and Geriatrics* 103 (Aug. 2022), 104795. https://doi.org/10.1016/j.archger.2022.104795

[17] Karin Hellfeldt, Laura López-Romero, and Henrik Andershed. 2019. Cyberbullying and Psychological Well-being in Young Adolescence: The Potential Protective Mediation Effects of Social Support from Family, Friends, and Teachers. *International Journal of Environmental Research and Public Health* 17, 1 (Dec. 2019). https://doi.org/10.3390/ijerph17010045

[18] Tom Hitron, Benny Megidish, Etay Todress, Noa Morag, and Hadas Erel. 2022. AI bias in Human-Robot Interaction: An evaluation of the Risk in Gender Biased Robots. In *2022 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 1598–1605. https://doi.org/10.1109/RO-MAN53752.2022.9900673

[19] Anna Lena Hunkenschroer and Christoph Luetge. 2022. Ethics of AI-Enabled Recruiting and Selection: A Review and Research Agenda. *Journal of Business Ethics* 178, 4 (Feb. 2022), 977–1007. https://doi.org/10.1007/s10551-022-05049-6

[20] Melissa Jauch, Selma Carolin Rudert, and Rainer Greifeneder. 2022. Social pain by non-social agents: Exclusion hurts and provokes punishment even if the excluding source is a computer. *Acta Psychologica* 230 (Sept. 2022), 103753. https://doi.org/10.1016/j.actpsy.2022.103753

[21] Mary E. Kite, Bernard E. Whitley, and Lisa S. Wagner. 2022. *Psychology of Prejudice and Discrimination* (4 ed.). Routledge, New York. https://doi.org/10.4324/9780367809218

[22] Alina Köchling and Marius Claus Wehner. 2020. Discriminated by an algorithm: a systematic review of discrimination and fairness by algorithmic decision-making in the context of HR recruitment and HR development. *Business Research* 13, 3 (Nov. 2020), 795–848. https://doi.org/10.1007/s40685-020-00134-w

[23] Ruth Levitas, Christina Pantazis, Eldin Fahmy, David Gordon, Eva Lloyed, and Demi Patsios. 2007. *The multi-dimensional analysis of social exclusion.* https://dera.ioe.ac.uk/id/eprint/6853/1/multidimensional.pdf

[24] Akansha Mer. 2023. Artificial Intelligence in Human Resource Management: Recent Trends and Research Agenda. In *Digital Transformation, Strategic Resilience, Cyber Security and Risk Management*, Simon Grima, Eleftherios Thalassinos, Grațiela Georgiana Noja, Theodore V. Stamataopoulos, Tatjana Vasiljeva, and Tatjana Volkova (Eds.). Vol. 111B. Emerald Publishing Limited, 31–56. https://doi.org/10.1108/S1569-37592023000111B003

[25] Ishaq Azhar Mohammed. 2021. Identity Management Capability Powered by Artificial Intelligence to Transform the Way User Access Privileges Are Managed, Monitored and Controlled. *International Journal of Creative Research Thoughts* 9, 1 (Jan. 2021), 4719–4723.

[26] Craig Morgan, Tom Burns, Ray Fitzpatrick, Vanessa Pinfold, and Stefan Priebe. 2007. Social exclusion and mental health: Conceptual and methodological review. *The British Journal of Psychiatry* 191 (Dec. 2007), 477–483. https://doi.org/10.1192/bjp.bp.106.034942

[27] Chloé Motillon-Toudic, Michel Walter, Monique Séguin, Jean-Daniel Carrier, Sofian Berrouiguet, and Christophe Lemey. 2022. Social isolation and suicide risk: Literature review and perspectives. *European Psychiatry* 65, 1 (Oct. 2022), e65. https://doi.org/10.1192/j.eurpsy.2022.2320

[28] Eirini Ntoutsi, Pavlos Fafalios, Ujwal Gadiraju, Vasileios Iosifidis, Wolfgang Nejdl, Maria-Esther Vidal, Salvatore Ruggieri, Franco Turini, Symeon Papadopoulos, Emmanouil Krasanakis, Ioannis Kompatsiaris, Katharina Kinder-Kurlanda, Claudia Wagner, Fariba Karimi, Miriam Fernandez, Harith Alani, Bettina Berendt, Tina Kruegel, Christian Heinze, Klaus Broelemann, Gjergji Kasneci, Thanassis Tiropanis, and Steffen Staab. 2020. Bias in data-driven artificial intelligence systems—An introductory survey. *WIREs Data Mining and Knowledge Discovery* 10, 3 (Feb. 2020). https://doi.org/10.1002/widm.1356

[29] Valerio Pellegrini, Valeria de Cristofaro, Marco Salvati, Mauro Giacomantonio, and Luigi Leone. 2021. Social Exclusion and Anti-Immigration Attitudes in Europe: The mediating role of Interpersonal Trust. *Social Indicators Research* 155, 2 (Jan. 2021), 697–724. https://doi.org/10.1007/s11205-021-02618-6

[30] Dana Pessach and Erez Shmueli. 2021. Improving fairness of artificial intelligence algorithms in Privileged-Group Selection Bias data settings. *Expert Systems with Applications* 185 (July 2021), 115667. https://doi.org/10.1016/j.eswa.2021.115667

[31] Florian Pethig and Julia Kroenung. 2023. Biased Humans, (Un)Biased Algorithms? *Journal of Business Ethics* 183, 3 (Feb. 2023), 637–652. https://doi.org/10.1007/s10551-022-05071-8

[32] Laura Pohlan. 2019. Unemployment and social exclusion. *Journal of Economic Behavior & Organization* 164 (Aug. 2019), 273–299. https://doi.org/10.1016/j.jebo.2019.06.006

[33] SAIL. 2023. SAIL - Sustainable Life-cycle of Intelligent Socio-Technical Systems. https://www.sail.nrw/

[34] Eric D. Wesselmann, Michelle R. Grzybowski, Diana M. Steakley-Freeman, Eros R. DeSouza, John B. Nezlek, and Kipling D. Williams. 2016. Social Exclusion in Everyday Life. In *Social Exclusion*, Paolo Riva and Jennifer Eck (Eds.). Springer, Cham, 3–23. https://doi.org/10.1007/978-3-319-33033-4_1

[35] Kipling D. Williams. 2009. Ostracism: A Temporal Need-Threat Model. In *Advances in Experimental Social Psychology*, Mark P. Zanna (Ed.). Vol. 41. Elsevier Academic Press, 275–314. https://doi.org/10.1016/S0065-2601(08)00406-1

[36] Lisa Zadro, Kipling D. Williams, and Rick Richardson. 2004. How low can you go? Ostracism by a computer is sufficient to lower self-reported levels of belonging, control, self-esteem, and meaningful existence. *Journal of Experimental Social Psychology* 40, 4 (July 2004), 560–567. https://doi.org/10.1016/j.jesp.2003.11.006

[37] Andrea Zechmann and Karsten Ingmar Paul. 2019. Why do individuals suffer during unemployment? Analyzing the role of deprived psychological needs in a six-wave longitudinal study. *Journal of Occupational Health Psychology* 24, 6 (Dec. 2019), 641–661. https://doi.org/10.1037/ocp0000154

Received 30 October 2023; accepted 20 November 2023