The Tension Between Information Justice and Security: Perceptions of Facial Recognition Targeting

Elizabeth Anne Watkins[a]

[a] *Center for Information Technology Policy, Princeton University, Princeton, NJ USA*

Abstract

In the discourse on human perceptions of algorithmic fairness, researchers have begun to analyze how these perceptions are shaped by sociotechnical context. In thinking through contexts of work, a half-century of research on organizational decision-making tells us that perceptions and interpretations made within these spaces are highly bounded by surrounding contextual constraints. In this paper I report early findings from a survey I conducted to bridge these two conversations, and scrutinize real-world perceptions of algorithmic decision-making *in situ* in a space of work. I analyze these perceptions through the case of facial recognition (or more accurately, facial verification) as account verification in gig work. In this survey I asked 100 Uber drivers, who all had been actually subjected to Uber's facial verification process known as Real Time Check ID, their fairness perceptions of this process. I designed the survey to elicit their perceptions across five disparate dimensions of justice: informational, distributive, procedural, reciprocal, and interactional. I also asked them about their strategies for integrating Real Time Check ID into their work flow, including efforts at repair when the system breaks down and their potential preferences for subversive practices. Of those workers who report engaging in subversive tactics to avoid facial recognition, such as taking a picture of their car seat, their hand, or their passenger instead of their own face, one dimension of fairness elicited worse perceptions than any other: informational justice, a.k.a. transparency, of facial recognition targeting (the process for deciding which workers trigger this extra layer of verification). This research reveals tensions between transparency, security, and workers' perceptions of the "fairness" of an algorithmic system: while "too much" transparency into how workers are targeted for verification may permit bad actors to defraud the system, "too little" explanation, this research shows, is no solution either. Results have crucial implications for the allocation of transparency and the design of explanations in user-facing algorithmic fraud detection, which must address tensions between information justice and security.

Keywords 1

Facial recognition, facial verification, biometric, algorithmic fairness, transparency, explainability, security, sociotechnical

1. Introduction

Questions around the fairness of algorithmic decision-making grow urgent as these systems proliferate across domains. "Fairness" suffers from what's called the "impossibility theorem," describing the multiplicity of fairness definitions — no fewer than 21 different definitions by one measure — and the impossibility of their simultaneous reconciliation [24]. The trouble is, machine learning success is contingent on building predictive accuracy, and predictive accuracy based on historical data is often antithetical to confirming human values. Values, after all, change over time, yet the definition, collection, and labeling of different metrics as required for machine learning elevate and bake into system design the values of specific stakeholders, in power at the time and place of design.

Joint Proceedings of the ACM IUI 2021 Workshops, April 13-17, 2021, College Station, USA
EMAIL: ew4582 [at] princeton.edu
ORCID: https://orcid.org/0000-0002-1434-589X
© 2020 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).
CEUR Workshop Proceedings (CEUR-WS.org)

CEUR-WS.org/Vol-2903/IUI21WS-TExSS-16.pdf

Prediction models provide binary outputs: in terms of facial identification for account verification in gig work, for example, the only outputs possible are either "the worker's face matches the photo on file," or "the worker's face doesn't match the photo on file," as delimited by a specified confidence threshold. At some point the developers of this model had to choose a threshold, to balance between too many false positives (where a low confidence threshold may yield cases where "this face matches" the profile picture, even in cases where it didn't), against false negatives (where an overly high threshold yields cases where "this face doesn't match" even in cases where it did). It's immediately apparent that when deciding where to place the confidence threshold, and which way to err in cases of error, high-power stakeholders like developers, executives, shareholders and even customers might have very different ideas of what constitutes a "fair" balance than workers, who have less say over system design.

A stream of research has begun to address how definitions of fairness are social constructs [4, 11], and the social, contextual conditions which contribute to the construction of fairness perceptions [17, 18, 35]. I contribute to this discourse by broadening the unit of analysis [2] of fairness perceptions, to incorporate this wider context in my study of gig workers and their interpretations of the fairness of facial recognition.[2] To bridge this work theoretically, I incorporate a theory of organizational decision-making known as the "fairness heuristic theory" [20].

Prior research on gig workers' online discussions of facial recognition illuminated potential for a rare empirical look at fairness interpretations *in situ*. In particular, qualitative analysis of Uber drivers' online discussions of facial recognition reveal that when they talk about their experiences with the technology, or "Real Time Check ID" as the protocol is called, they discuss — some obliquely, some directly — terms related to "fairness." This is especially apparent among workers who disagree with the requirement. In one salient example, one driver posted to an online discussion board the question, "How can I beat facial recognition?" When another driver replied, "Why would you need to beat facial recognition?" the first responded, "Any tactic used by a driver to combat the crooked set up is fair game" [34].

To shed light on the entanglement between perceptions of algorithmic fairness, sociotechnical context, and organizational decision-making, I designed a survey to illuminate relationships between how workers perceive the fairness of Real Time Check ID, across two involved processes: facial recognition targeting, i.e. the process deciding who gets chosen to comply with the protocol, and verification, i.e. the actual process using biometric data to verify identity. The goal of this survey is to gather fairness perceptions across a set of dimensions drawn from literature in human perceptions of fairness and organizational justice, as well as workers' practices, negotiations, and tactics around this technology. In this paper, I share descriptive findings comparing workers' fairness perceptions and their behaviors, with a focus on perceptions of information justice, i.e. "explanations provided to people that convey information about why procedures were used" to arrive at a decision [8]. Once we have data around perceptions and behaviors, we can begin to explore the relationships between respective perceptions, and identify whether specific perceptions have explanatory power for workers' behaviors.

2. Human Perceptions of Algorithmic Fairness in Sociotechnical Context

An emerging body of research is beginning to uncover how social context and human perceptions of algorithmic fairness interact. The interwoven nature of perceptions, behaviors, and structures which either constrain or promote these cognitive and functional processes contribute to the messy realities of institutional life for both algorithms and users. These studies have been essential to establish that users' ideas of fairness are highly contextual social constructs [35] and that even within equitable contexts algorithms can be

[2] Facial recognition is a general term which can apply to both facial verification (1:1 matching) as well as identification (1:n matching). The protocol I analyze in this research is more accurately described as facial verification (i.e. 1:1 matching, verifying that a face in an image uploaded from a phone matches the face in the driver's profile photo). Because I queried drivers about two distinct algorithmic protocols involving verification as well as targeting I use "facial recognition" as an umbrella term for both processes. I also use "selfie verification" as it's colloquially called by drivers.

subject to different interpretations [4]. Workers have little control over the algorithms that manage their work, and their perceptions of the fairness of algorithms within a work context are influenced by the relative control they're granted over the algorithmic decision [16]. Beunza [3] suggests that, when workers are directed by an algorithm that they perceive as unfair, this may increase their willingness to engage in unethical behavior. Several researchers have drawn important frameworks from management science literature, in particular the concept of organizational justice [11]. Organizational justice offers useful handles for analyzing multiple dimensions of how a worker is treated by a single algorithm within an organization [17]. One key insight is that users perceive algorithmic fairness along multiple dimensions, including: 1) distributive justice, i.e. fairness in the distribution of outcomes, 2) procedural justice, i.e. fairness of the logic underpinning a decision, 3) informational justice, i.e. understanding how decisions are made, and 4) interactional justice, i.e. whether participants feel decision-making processes treat them with dignity and respect [4].

When it comes to operationalizing context for research purposes, participants in such studies have most often been presented with speculative scenarios on which to base their perceptions. In available survey and participatory workshop studies, participants have been given a fictional second-person scenario, such as, "imagine you are applying for a personal financial loan," or "promotion at work," or "car insurance premiums," [4]. Some studies go further, and recruit participants from marginalized communities, and ask them how they feel about the types of contexts in which algorithmic decisions often take place, again via speculative scenarios (i.e. discriminatory advertising) [35]. This is where my research picks up the torch and contributes to the discourse on algorithmic fairness in sociotechnical context: by investigating fairness perceptions *in situ*, in real contexts with people actually engaged in and subject to the actual algorithms about which they're queried. In this research, I focus on the context of work.

3. Decision-Making in Organizations

Organizational theorists have long posited that decision-making in organizations is a messy process, heavily contingent on local interpretations, values, and available knowledge. Decisions around technology in organizations are no exception. Theorists in this school have drawn from language in Science, Technology, and Society (STS) Studies, to better understand how the social and the technical are co-constitutive in spaces of work [26]. Local practices and interpretations are influential components of how new technology becomes integrated into organizations [1, 25]. Workers alter their use of technologies depending on what they believe the tools are for, and resist prescriptive uses (i.e. directives from the organization) if those directives don't match their interpretations [19]. The power of interpretation can also be seen through the multiple accounts of friction and resentment created when workers' interpretations of a technology's use don't match those of the organization [7, 16]. The theory of bounded rationality, from mid-century organizational theory, illuminates how risk and uncertainty prevent anyone in an organization from making perfectly rational decisions [32]. Organizational inputs and outputs are unpredictable, and information is never perfect and often incomplete.

How do organizational members make choices in such a chaotic environment? "Fairness heuristic theory" [20] argues that people working in an organization address the cognitive load of decision-making, in particular tensions between their individual autonomy and their group identity, by drawing on a cognitive shortcut of fairness derived from their perceptions of how fairly they've been treated in other decisions in the organization. The fairness heuristic provides a framework to scrutinize users' perceptions in AI-infused sociotechnical systems of work, and the relationship between these perceptions and subsequent behaviors. My first step is to assess fairness perceptions across multiple dimensions and organizational processes, to establish a basis of comparison. This research embraces a sociotechnical lens on algorithmic fairness, recognizing the intertwined influence of the social and technical components of a system

[9]. What a technology means, how it's interpreted and communicated, and what problems it addresses, arise from the interplay of both social and technical aspects of a system. Without attention to the local, contingent nature of knowledge in sites of integration, the design of algorithmic systems is likely to suffer from a number of abstractions which may eventually damage actors in that system [30].

4. Background

Facial recognition as a form of account security has been used by Uber on drivers under the name "Real Time Check ID" since it was rolled out in select regions in 2016.[3] This involves two primary interwoven algorithmic processes: 1) targeting, which uses behavioral data and algorithmic modeling to detect potentially fraudulent activity on the app and select drivers whose behavior is flagged for additional checks, and 2) verification, the algorithmic processes which verify, or confirm, that the face of the person behind the wheel and logged into the platform matches the photo on the profile, of the person who's been approved to work by Uber.

What does this look like in practice? When a driver logs onto the app, a circle pops up on the screen with a text command to the driver to position their face inside the circle. A photo is taken using the driver's phone camera. That photo (or a data-based representation of that photo) is sent to Microsoft, where computer-vision software-as-a-service compares it with the "official" photo of the driver on the account. If it is decided that the faces "match," the driver is considered "verified," and allowed to log into the platform and start working. If it is decided that the faces "don't match," the driver must take another photo. They are not permitted to log into the platform until a match is reached. They must either continue to re-take photos until they get verified, or, if verification continues to fail, go to a customer-service center in person to get assistance.

5. Survey Methods

This descriptive, exploratory survey measures and compare workers' perceptions of five types of fairness across of Real Time Check ID targeting and verification. I adapted fairness perception questions from recent work [4] drawing on the psychology of justice research, identifying the multiple, simultaneous parameters of fairness perception: *Procedural* justice concerns the processes of making a decision. *Distributive* justice concerns how results are allocated across participants or between groups, also known in some circles as outcome fairness. *Interactional* justice concerns the extent to which the affected individual is treated with respect by the decision-makers. *Informational* justice pertains to the information available to participants about the decision-making process. *Reciprocal* justice pertains to individuals' comparisons between their inputs versus their outputs in their involvement with an exchange, and whether they're getting a fair return for their efforts. The last dimension, of Reciprocal justice, was motivated by literature examining perceptions of fairness within organizations, in particular around equity [13]. I added this dimension to capture perceptions related to the labor that users put into selfie verification in exchange for access to the platform, and to recognize that workers do not passively receive an algorithmic judgment but rather are active creators and maintainers of the conditions that allow algorithms to make that judgment.

All five of these fairness dimensions were worded into statements for each decision-making process, for which participants could select a response along a five-point scale of "Strongly Agree," "Somewhat Agree," "Neither Agree nor Disagree," "Somewhat Disagree," or "Strongly Disagree," as follows:

Real Time Check ID Targeting:

Distributive: The process for deciding who gets chosen for selfie verification treats all drivers equally.

Informational: I understand how drivers are chosen for selfie verification.

Interactional: The process for how drivers are chosen for selfie verification shows respect for me and the work that I do.

Reciprocal: The benefits that I receive from selfie verification are fair, as compared with the time and effort I spend operating and/or trouble-shooting.

[3] Implementation of this protocol varies in accordance with local data protection legislation. In the U.K., for example, with GDPR data protections, drivers can request human review instead of algorithmic review.

Procedural: The way Uber decides which drivers get the pop-up for selfie verification is a fair way to decide who has to verify their account.

Real Time Check ID Verification:

Distributive: Selfie verification treats all drivers equally.

Informational: I understand how selfie verification works to verify my identity.

Interactional: Selfie verification shows respect for me and the work that I do.

Reciprocal: The benefits that I receive from selfie verification are fair, as compared with the time and effort I spend operating and/or trouble-shooting.

Procedural: Using selfie verification to verify a driver's identity is a fair way to ensure that drivers' accounts are secure.

Participants for this survey were recruited and participated in this survey using the platform Qualtrics. A number of platforms exist for researchers to recruit survey participants, all of which feature different benefits and drawbacks. While some findings vary as to which web-based survey platforms yield results that are representationally similar to the U.S. population, others have found that Qualtrics' participant population is the most demographically and politically representative [5]. Because demographics and racial justice have been so important to recent studies in algorithmic bias, in particular facial recognition, Qualtrics was selected as the platform both for recruitment and for the survey itself. All participants had previously signed up to be available potential participants on Qualtrics. Typically, respondents choose to join a survey through a double opt-in process. Upon registration, they enter basic data about themselves, including demographic information, education, job title, hobbies, and interests. Whenever a survey is created that that individual would qualify for based on the information they have given, they are notified via email and invited to participate. The email invitation is generic, with no specifics as to the topic of the survey itself. They are told that they qualify for a survey, given a link, and asked to follow the link if they would like to participate. They are also told the duration of the survey. Participants in this survey were compensated for their time and effort on a points system, redeemable towards air miles and gift cards. Recruitment quotas for race, ethnicity, and self-identified gender were established using the most recent available data on Uber's labor market in the United States [12] to gather a population resembling, as much as possible, the U.S. Uber labor market.

6. Results

One hundred participants took part in this survey over the course of five days in July 2020. This is a small dataset, and is not intended to be taken as statistically representative of the entirety of the U.S. Uber labor market. However, this data is proportionally similar to that market, can provide some descriptive idea of the perceptions of drivers in a group, and can lend clarity to potential next steps in research. In accordance with survey practices set by Pew and others, race and ethnicity were broken into separate questions for the survey, with Hispanic ethnicity in a separate question from that of race identity. These choices were then re-condensed for analysis purposes to assure that the final dataset resembled available information on the Uber U.S. labor market. All group representation resembles their representation in the United States Uber driver population along lines of race and gender identity to within three percentage points. The use of screening questions ensured that 100% of participants self-report currently driving for Uber and that they have actually been required at least once to comply with Real Time Check ID. 100% are within the United States to assure proportionality to available statistics on Uber's labor market. The survey totaled 34 questions covering drivers' experiences and perceptions, with some demographic questions. Average time to complete the survey was 3.05 minutes, with a median time of 4.03 minutes. Seven percent of participants were under the age of 25, 34% were 25-34, 48% were 35-44, seven percent were 45-54, and four percent were 66-64.

In their responses to fairness statements, one type of statement provoked a far lower rate of agreement than any other statement across any dimension or process (see **Figure 1**): drivers disagree most with the statement on information justice of Real Time Check ID targeting. This fairness statement was phrased as "I understand how drivers are chosen for selfie verification," a colloquial way of describing a state in which drivers are provided with enough information to understand how targeting decisions are made. Put simply, drivers disagree with the idea that targeting decisions made about them are transparent or understandable.

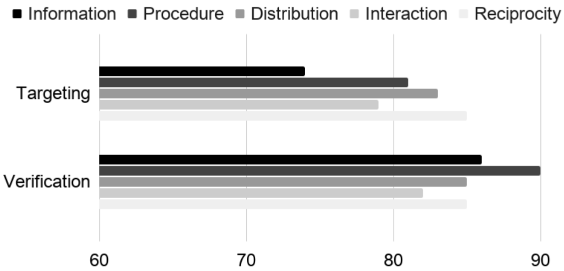

Figure 1: Drivers' total rates of agreement with fairness statements about Real Time Check ID targeting (top) and verification (bottom) across five dimensions of justice.

What relationship, if any, can be seen between this perception and drivers' behaviors? To gather information about behaviors, I asked a subset of drivers (n=75) about their Real Time Check ID compliance preferences, with the question "In the event that you PREFER not to comply with selfie verification, which tactics have you used?" Nearly two-thirds (65.3%) responded that they had used subversive tactics to avoid showing their face to the camera. Over one-third (38.7 percent) reported that they had submitted, instead of their face, "an unusual photo with selfie verification hoping the system will approve it (like of the car seat, or the passenger)."

I then looked specifically at this group's information justice perceptions. Among these "subversive" drivers, rates of agreement with information justice statements about targeting and verification drop, across the board (see **Figure 2**).

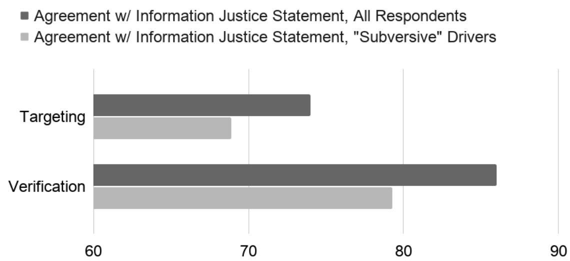

Figure 2: Drivers' total rates of agreement specifically with information justice statements about Real Time Check ID targeting (top) and verification (bottom) for all respondents (dark grey) and "subversive" drivers, i.e. those who report they prefer not to comply (light grey).

Among the entire set of drivers, the total rate of agreement around the information justice of Real Time Check ID targeting is 74%; among subversive drivers, that rate drops to 68.9%. Among all drivers, the total rate of agreement around the information justice of verification is 86%. Among subversive drivers, that rate drops to 79.3%.

7. Conclusion and Discussion

This research describes three findings regarding drivers' perceptions and behaviors around the fairness of algorithmic security protocols in their work. First, among five justice statements around the protocols of facial recognition (Real Time Check ID targeting and verification), the lowest rates of driver agreement concern the information justice of Real Time Check ID targeting. Drivers disagree at comparatively high rates with statements that they understand how they're being targeted for selfie verification.

Second, this research yields potential evidence for an observable relationship between perceptions and their subsequent behaviors. Among workers who self-report that they prefer not to comply with facial recognition protocols (and act on that preference), there is a marked drop in agreement with information justice statements regarding targeting, as compared to the entire group of participants.

Third, within this group, similar drops are observable across two disparate algorithmic processes: Real Time Check ID targeting as well as Real Time Check ID verification. More research is needed to better understand whether such similarities constitute a fairness heuristic. Such research may investigate for example whether drivers have similar fairness perceptions across different processes of algorithmic management such as task allocation and pricing, or whether fairness perceptions of algorithmic processes have any relationship with subsequent behaviors for other processes related to that workflow or work platform.

It's difficult to overstate how much gig work platforms use information asymmetries and algorithmic mechanisms in order to re-allocate risk, uncertainty, and ambiguity onto workers [21, 27, 28, 22]. The passage of Proposition 22 in California has in particular cemented the definition of gig work there as contract work without employer protections. Research about "fairness" perceptions of algorithmic protocols in such contexts must recognize these asymmetries, and further, how the availability (or lack thereof) of other options may influence how "fair" workers perceive their algorithmic options to be. Future research could explore how precarity and economic reliance contribute to perceptions of transparency, fairness, and justice.

A wealth of research has demonstrated how recognition technologies are built on extractive means for carceral ends, using surveillance technologies to build ever-larger datasets which met out harms due to the faulty functionality of the technology, largely for Black, Indigenous, and people of color [6, 28]. The market for emotion recognition, for example, continues to grow, despite the fact that the scientific foundation for such recognition claims are dubious at best and violate human rights at worst [23]. Recognition technologies also produce violence for groups such as trans and nonbinary people through the functional design of categorization, which is a poor match for the recognition of human identity which is fluid and malleable [15, 31]. Black and transgender workers have already been locked out of Uber due to Real Time Check ID [14, 33]. Recent research has brought about critically needed legislative responses to protect people. More research is needed into how the intrinsic injustices of recognition technology intersect with the injustices of algorithmic management and precarious work to produce distributed harms.

Specifically, these findings provoke questions about the relationships between algorithmic transparency, information justice, and effective security protocols. A tension emerges between informational justice for drivers and the need to obscure the details of how Real Time Check ID targeting works. If targeting algorithms processes and factors were made more transparent to drivers, this might become a vulnerability which bad actors could exploit. "Too much" transparency into how drivers are targeted for verification may permit bad actors to defraud the system. Yet, "too little" explanation, this research shows, is no solution either: a lack of information justice (what we could call, information injustice) seems to correlate with "subversive" practices, which may be categorized as deviance and result in drivers being barred from the platform.

These results have crucial implications for the design of explanations and transparency in user-facing algorithmic fraud detection, which must address tensions between informational justice and security.

Acknowledgments

This work was made possible by support provided by the UC Berkeley Center for Long-Term Cybersecurity, for their symposium Data Rights, Privacy, and Security in AI. Thanks to the participants of that symposium, especially Shazeda Ahmed, Richard Wong, and Karen Levy. Support was also provided via a travel grant from the ACM Conference on Fairness, Accountability, and Transparency (FAccT). Invaluable feedback was provided by Susan McGregor and Amanda Lenhart. This work would not be possible without the guidance of Kelly Caine and the feedback of her graduate students at Clemson University. Additional thanks to the AI on the Ground group at the Data & Society Research Institute, including Madeleine Elish, Emanuel Moss, Jake Metcalf, and Ranjit Singh, and the CITP group at Princeton, especially Amy Winecoff. Lastly, thanks to the anonymous reviewers' generous and thoughtful feedback.

References

[1] Bailey, D.E. and Leonardi, P.M., 2015. Technology choices: Why occupations differ in their embrace of new technology. MIT Press

[2] Barad, K., 2014. Diffracting diffraction: Cutting together-apart. Parallax, 20(3), pp.168-187.
[3] Beunza, Daniel. 2019. Taking the Floor: Models, Morals, and Management in a Wall Street Trading Room. Princeton: Princeton University Press.
[4] Binns, R., Van Kleek, M., Veale, M., Lyngs, U., Zhao, J. and Shadbolt, N., 2018, April. 'It's Reducing a Human Being to a Percentage' Perceptions of Justice in Algorithmic Decisions. In Proceedings of the 2018 Chi conference on human factors in computing systems (pp. 1-14).
[5] Boas, T.C., Christenson, D.P. and Glick, D.M., 2020. Recruiting large online samples in the United States and India: Facebook, mechanical turk, and qualtrics. Political Science Research and Methods, 8(2), pp.232-250.
[6] Buolamwini, J. and Gebru, T., 2018, January. Gender shades: Intersectional accuracy disparities in commercial gender classification. In Conference on fairness, accountability and transparency (pp. 77-91).
[7] Christin, A., 2017. Algorithms in practice: Comparing web journalism and criminal justice. Big Data & Society, 4(2), p.2053951717718855.
[8] Colquitt, J. A., Conlon, D. E., Wesson, M. J., Porter, C. O. L. H., & Ng, K. Y. 2001. Justice at the millennium: A meta-analytic review of 25 years of organizational justice research. Journal of Applied Psychology, 86, 425-445.
[9] Elish, Madeleine Clare and Elizabeth Anne Watkins. 2020. Repairing Innovation: A Study of Integrating AI in Clinical Care (New York: Data & Society Research Institute)
[10] Griesbach, K., Reich, A., Elliott-Negri, L. and Milkman, R., 2019. Algorithmic control in platform food delivery work. *Socius*, *5*, p.2378023119870041.
[11] Grgic-Hlaca, Nina, Elissa M. Redmiles, Krishna P. Gummadi, and Adrian Weller. 2018. "Human perceptions of fairness in algorithmic decision making: A case study of criminal risk prediction." In Proceedings of the 2018 World Wide Web Conference, pp. 903-912.
[12] Hall, J. and Alan Krueger. 2016. An Analysis of the Labor Market for Uber's Driver-Partners in the United States. National Bureau of Economic Research Working Paper Series. Accessed at https://www.nber.org/papers/w22843.pdf
[13] Joshi, K. 1989. The Measurement of Fairness or Equity Perceptions of Management Information Systems Users. MIS Quarterly , Sep., 1989, Vol. 13, No. 3 (Sep., 1989), pp. 343-358
[14] Kersley, Andrew. 2021. WIRED. "Couriers say Uber's 'racist' facial identification tech got them fired." https://www.wired.co.uk/article/uber-eats-couriers-facial-recognition
[15] Keyes, O., 2018. The misgendering machines: Trans/HCI implications of automatic gender recognition. Proceedings of the ACM on Human-Computer Interaction, 2(CSCW), pp.1-22.
[16] Lebovitz, S., Levina, N. and Lifshitz-Assaf, H., 2019. Doubting the diagnosis: How artificial intelligence increases ambiguity during professional decision making. Available at SSRN 3480593.
[17] Lee, M.K., Jain, A., Cha, H.J., Ojha, S. and Kusbit, D., 2019. Procedural justice in algorithmic fairness: Leveraging transparency and outcome control for fair algorithmic mediation. Proceedings of the ACM on Human-Computer Interaction. 3(CSCW), pp.1-26.
[18] Lee, M.K., Kusbit, D., Metsky, E. and Dabbish, L., 2015, April. Working with machines: The impact of algorithmic and data-driven management on human workers. In Proceedings of the 33rd annual ACM conference on human factors in computing systems (pp. 1603-1612).
[19] Leonardi, P.M., 2009. Why do people reject new technologies and stymie organizational changes of which they are in favor? Exploring misalignments between social interactions and materiality. Human Communication Research, 35(3), pp.407-441.
[20] Lind, E.A., 2001. Fairness heuristic theory: Justice judgments as pivotal cognitions in organizational relations. Advances in organizational justice, 56(8).
[21] Moradi, P. and Levy, K., 2020. The Future of Work in the Age of AI: Displacement or Risk-Shifting?. The Oxford Handbook of Ethics of AI. pp. 271-87 (Markus Dubber, Frank Pasquale, and Sunit Das, eds.)

[22] Moss, E. and Metcalf, J., 2020. High Tech, High Risk: Tech Ethics Lessons for the COVID-19 Pandemic Response. *Patterns*, *1*(7), p.100102.

[23] Marda, Vidushi and Shazeda Ahmed. 2021. ”Emotional Entanglement: China’s emotion recognition market and its implications for human rights.” Article 19.

[24] Narayanan, A., 2018, February. Translation tutorial: 21 fairness definitions and their politics. In Proc. Conf. Fairness Accountability Transp., New York, USA (Vol. 1170).

[25] Orlikowski, W.J. and Gash, D.C., 1994. Technological frames: making sense of information technology in organizations. ACM Transactions on Information Systems, 12(2), pp.174-207.

[26] Orlikowski, W.J. and Scott, S.V., 2008. 10 sociomateriality: challenging the separation of technology, work and organization. Academy of Management annals, 2(1), pp.433-474.

[27] Qadri, R., 2020, February. Algorithmized but not Atomized? How Digital Platforms Engender New Forms of Worker Solidarity in Jakarta. In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society* (pp. 144-144).

[28] Raji, I.D. and Buolamwini, J., 2019, January. Actionable auditing: Investigating the impact of publicly naming biased performance results of commercial ai products. In *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society* (pp. 429-435)

[29] Rosenblat, A. and Stark, L., 2016. Algorithmic labor and information asymmetries: A case study of Uber’s drivers. *International journal of communication*, *10*, p.27.

[30] Selbst, A.D., boyd, D., Friedler, S.A., Venkatasubramanian, S. and Vertesi, J., 2019, January. Fairness and abstraction in sociotechnical systems. In Proceedings of the Conference on Fairness, Accountability, and Transparency (pp. 59-68).

[31] Scheuerman, M.K., Paul, J.M. and Brubaker, J.R., 2019. How computers see gender: An evaluation of gender classification in commercial facial analysis services. Proceedings of the ACM on Human-Computer Interaction, 3(CSCW), pp.1-33.

[32] Simon, H., 1957. A behavioral model of rational choice. Models of man, social and rational: Mathematical essays on rational human behavior in a social setting, pp.241-260.

[33] Urbi, Jaden. 2018. "Some transgender drivers are being kicked off Uber’s app." /www. cnbc.com/2018/08/08/transgender-uber-driver-suspended-tech-oversight-facial-recognition.html

[34] Watkins, E.A., 2020, October. Took a Pic and Got Declined, Vexed and Perplexed: Facial Recognition in Algorithmic Management. In Conference Companion Publication of the 2020 on Computer Supported Cooperative Work and Social Computing (pp. 177-182).

[35] Woodruff, A., Fox, S.E., Rousso-Schindler, S. and Warshaw, J., 2018, April. A qualitative exploration of perceptions of algorithmic fairness. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (pp. 1-14)