Association for Information Systems
AIS Electronic Library (AISeL)

SIGHCI 2017 Proceedings

Special Interest Group on Human-Computer Interaction

12-2017

Demographic Transparency to Combat Data Analytics Discriminatory Recommendations

Sepideh Ebrahimi
McMaster University, s.ebrahimi@mcmaster.ca

Khaled Hassanein
McMaster University, hassank@mcmaster.ca

Demographic Transparency to Combat Data Analytics Discriminatory Recommendations

Sepideh Ebrahimi
DeGroote School of Business
McMaster University
s.ebrahimi@mcmaster.ca

Khaled Hassanein
DeGroote School of Business
McMaster University
hassank@mcmaster.ca

ABSTRACT

Data Analytics (DA) has been blamed for contributing to discriminatory managerial decisions in organizations. To date, most studies have focused on the technical antecedents of such discriminations. As a result, little is known about how to ameliorate the problem by focusing on the human aspects of decision making when using DA in organizational settings. This study represents an effort to address this gap. Drawing on the cognitive elaboration model of ethical decision-making, construal level theory, and the literature on moral intensity, this study investigates how the availability and the design of demographic transparency (a form of decisional guidance) can lower DA users' likelihood of agreement with discriminatory recommendations of DA tools. In addition, this study examines the role of user's mindfulness and organizational ethical culture on this process. This paper outlines an experimental methodology to empirically validate the proposed model and hypotheses and delineates potential contributions to theory and practice.

Keywords

Data Analytics, Discrimination, Demographic Transparency, Ethical Decision Making, Proximity.

INTRODUCTION

The last decade, technological advances have enabled organizations to collect an ever-increasing amount of data, which they strive to analyze by employing data analytics (DA) tools in order to make data-driven decisions. Such an approach to decision making, though suggested to bring about several benefits, has been accused of contributing to discrimination in societies (Newell and Marabelli, 2015).

Discriminatory recommendations of DA tools are mainly generated due to abundance of data on individuals being analyzed, biased or non-representative data, and inadvertent modeling procedures (Žliobaitė and Custers 2016). While some technical methods have been suggested to alleviate the issue (e.g., Pedreshi, Ruggieri and Turini, 2008), developing computational means to prevent such discrimination is still an ongoing endeavor. The insufficiency of the existing methods to eliminate the issue is evident in recent scholarly and practitioners' (Crawford, 2013; Newell and Marabelli, 2015) and even governments' (Podesta, Pritzker, Moniz, Holdren and Zients, 2014) publications that raise concerns about the potential of discriminatory recommendations m DA tools.

In organizations, it is ultimately the decision makers' responsibility to make sure that their data-driven decisions are free of discrimination. However, Newell and Marabelli (2015) suggest that few individuals actually understand what is included in the algorithms and why. Therefore, this study focuses on the outcomes of DA systems to help DA users reduce the instances of discriminatory decisions. More specifically, this study aims at investigating *how and to what extent providing aggregated demographic information regarding the human subjects of the DA recommendations would reduce the incidence of users' agreement with discriminatory recommendations of DA systems?*

The role of individual and organizational characteristics in making (un)ethical decisions has been vastly studied. One such individual characteristic is mindfulness, which has received attention in both the IS and ethics literatures. In addition, the important role of an organizational ethical culture in influencing individuals' beliefs and behavior has been underlined in many studies (Hunt and Vitell, 1986). Therefore, this study also strives to explore *how and to what extent would organizational ethical culture and user's mindfulness impact the relationship between providing the aforementioned aggregated demographic information and data analytics users' agreement with discriminatory recommendations of those systems?*

THEORETICAL BACKGROUND

Cognitive elaboration model of ethical decision-making

During the course of making an ethical decision, individuals move through a series of four steps: recognition of the moral issue, making a moral judgment, establishing the intent to act morally, and engaging in a moral behavior (Rest, 1986). The first step, recognizing the moral aspect, is a form of attitude change toward an object/event (Street, Douglas, Geiger and Martinko, 2001). As such, according to Elaboration Likelihood Model, when an individual faces an ethically charged issue, if the overall influence of the available ability and motivation factors results in a high level of elaboration,

the decision maker uses their central information processing route and, thus, is more likely to recognize the moral aspects of the issue at hand (Street et al., 2001).

Moral Intensity (MI)

One motivational variable discussed by Street et al. (2001) is the notion of moral intensity (MI) introduced by Jones (1991). He suggests moral intensity of an issue to be comprised of six factors: magnitude of consequences, social consensus, probability of effect, temporal immediacy, proximity, and concentration of effect (Jones, 1991). These dimensions of MI, have been widely shown to have a high impact on the ethical decision making process (e.g., Singhapakdi, Vitell and Kraft, 1996). Our focus in this study is on increasing DA users' perceived proximity toward the subjects of their decision. This is because computers in general and more specifically DA tools tend to distance the decision maker from the subjects of their decisions (Ebrahimi, Ghasemaghaei and Hassanein, 2016) and the further the perceived distance of the victims of the action to the decision maker, the less intense will be the moral issue in their mind. The notion of distance has a high interrelationship with the notion of construal level as discussed by the construal level theory.

Construal Level Theory (CLT)

CLT (Trope and Liberman, 2010) suggests that people use a more abstract, high construal level when perceiving and predicting more psychologically distal targets, and they judge more abstract targets as being more psychologically distal (Bar-Anan, Liberman and Trope, 2006). Psychological distance is the "subjective experience that something is close or far away from the self, here, and now" (Trope and Liberman, 2010, p. 440). Trope and Liberman (2010) discuss the relationship between psychological distance and construal level and suggest that because less information is available about distal entities than proximal ones, people typically construe the former more abstractly than the latter. In addition, since high-level construals are more general, they tend to bring to mind more distal instantiations of the entities. Therefore, an association is made between the psychological distance and construal level.

RESEARCH MODEL AND HYPOTHESES DEVELOPMENT

To respond to the two research questions outlined earlier, this study draws upon the above theoretical foundations, to propose the research model (depicted in Figure 1) and associated hypotheses, detailed below with support.

Demographic Transparency (DT)

As previously discussed, discriminatory recommendations by DA tools often arise due to reasons other than basing the recommendation on sensitive variables and therefore are instances of indirect discrimination. While, to the best of our knowledge, no rigid mathematical formula exists to frame indirect discrimination, there have been some suggestions as to what constitutes an instance of it. For example, "the four-fifth rule" adopted by several American Institutes (e.g., Department of Labor), states that a selection rate for any race, sex, or ethnic group, which is less than four-fifths of the selection rate for the group with the highest selection rate will be regarded as evidence of disparate impact (Feldman, Friedler, Moeller, Scheidegger and Venkatasubramanian, 2015). Due to the common practice of identifying discriminatory decisions through investigating their outcomes and the complexity of investigating the process of a discriminatory recommendation being put forth by DA tools, this study focuses on the outcomes of DA tools analyses and suggests providing DA users with aggregated demographic information about the data subjects.

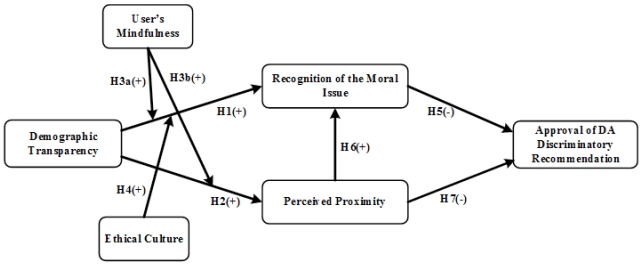

Figure 1. Proposed Research Model

Such aggregated information is a form of decisional guidance defined by Silver (1991, p. 107) as "how a decision support system enlightens or sways its users as they structure and execute their decision-making processes". The particular form of decisional guidance to be examined in this study is *demographic transparency (DT),* which is defined as *informative* guidance aimed at helping users compare the proportion of each demographic class (e.g., female and male) in its pertinent demographic category (e.g., gender) in the original full dataset and the DA-recommended sample. For instance, as Figure 2 depicts, imagine a situation, where the original pool includes 1000 applicants for a position from which 44% are female and 56% are male. The system's recommendation will be judged to be potentially discriminatory if its recommended sample of applicants to be considered for the position includes 21% and 79% females and males respectively as the proportion of male

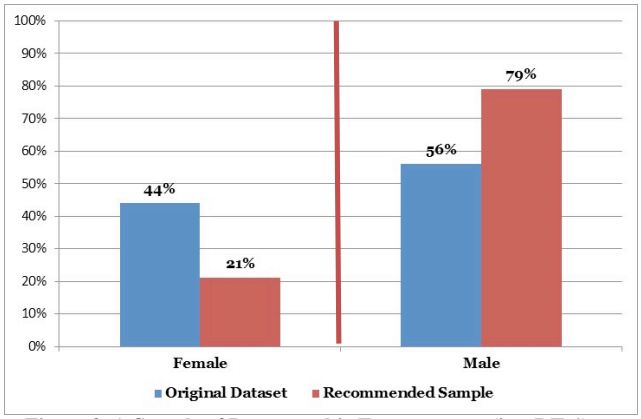

Figure 2. A Sample of Demographic Transparency (i.e., DT-1)

(female) applicants who were included in the recommended sample is considerably more (less) than the proportion of male (female) applicants in the original pool. This study suggests that providing DA users with DT increases their ability to recognize traces of potential discrimination (if any) in the system's recommendations. The user can then gauge whether discrimination exists based on the context of the decision.

In addition to increasing users' ability, we suggest that providing DT decreases the user's mental construal level of and the proximity felt toward the subjects of their decision. This is because high-level construals are viewed as "relatively abstract, coherent, and superordinate mental representations, compared with low-level construals" (Trope and Liberman, 2010, p. 441). Therefore, the less abstract and congruent a group of people are perceived to be by an individual, the higher the chance that he/she will have a lower level construal of them.

To further increase the proximity that a DA user feels toward the subjects of their decision, this study suggests increasing the level of DT by adding representative images of each demographic class next to its pertinent information on the chart (Figure 3). As suggested by CLT, pictures bear physical resemblance to the referent objects, whereas words are abstract representations that carry the essence of that object (Amit, Algom and Trope, 2009). Hence, pictures comprise a lower level of construal than do words and therefore, are more likely to result in a feeling of proximity to the receiver. Therefore,

H1: *In the context of a potential DA discriminatory recommendation, DA tools with higher demographic transparency will increase the likelihood of users' recognition of the moral issue.*

H2: *In the context of a potential DA discriminatory recommendation, DA tools with higher demographic transparency will increase users' perceptions of proximity toward the subjects of their decision.*

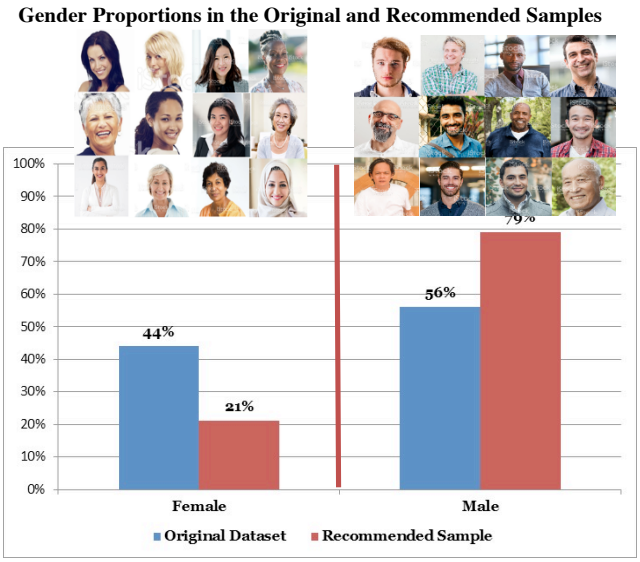

Figure 3. A Sample of Demographic Transparency with Representative Images (i.e., DT-2)

User's Mindfulness

Many unethical decisions stem from a lack of awareness and moral awareness is the critical first step in the ethical decision making process (Rest, 1986). Mindfulness is a notion that goes hand-in-hand with awareness and is defined as "an enhanced attention to and awareness of current experience or present reality" (Brown and Ryan, 2003, p. 822). Mindfulness has been shown to have an inverse relationship with willingness to engage in unethical behavior (Ruedy and Schweitzer, 2010). This study suggests that the relationships between DT and recognition of the moral issue and proximity are positively moderated by users' mindfulness. This is because mindfulness captures a quality of consciousness that is mainly thought of as the vividness of one's current experience and hence, is in contrast to the mindless or automatic functioning (Brown and Ryan, 2003). Thus,

H3: *In the context of a potential DA discriminatory recommendation, users' mindfulness moderates the relationships between demographic transparency and (a) recognition of the moral issue; (b) perceived proximity toward the subjects of their decision, such that the effects are stronger for individuals higher in mindfulness than for those lower in mindfulness.*

Ethical Culture

Ethical culture refers to "a subset of organizational culture, representing a multidimensional interplay among various "formal" and "informal" systems of behavioral control that are capable of promoting either ethical or unethical behavior" (Treviño, Butterfield and McCabe, 1998, p. 451). The positive impact of the organization's ethical culture on its employees' moral awareness and behavior has been demonstrated in the literature (Craft, 2013). Indeed, culture influences individuals' beliefs (Hunt and Vitell, 1986). From a cognitive perspective, particular thought processes can be invoked by individual's exposure to cultural cues (Hong, Morris, Chiu and Benet-Martinez, 2000). Hence, we suggest that in an organization where salient values prompt ethical thinking, it is more likely that ethical thought processes are invoked in employees' minds as a result of being exposed to DT. Therefore,

H4: *In the context of a potential DA discriminatory recommendation, ethical culture of the organization moderates the relationships between demographic transparency and recognition of the moral issue, such that the effect is stronger for individuals from organizations with stronger ethical cultures.*

Recognition of the Moral Issue, Proximity, and Approval of the Discriminatory Recommendation

Ethical reasoning has been described as a systematic framework. Individuals first realize an ethical situation, which prompts them to consider and evaluate courses of

actions based on their morality. Such assessments subsequently affect their ethical intentions and action (Rest, 1986). Many studies to date have found significant relationships between the aforementioned stages (For a review, see Craft, 2013). Therefore,

H5: *In the context of a potential DA discriminatory recommendation, users' recognition of the moral issue is negatively associated with their approval of the discriminatory recommendation.*

Proximity, a dimension of MI, is defined as the "feeling of nearness that the moral agent has for victims (beneficiaries) of the detrimental (beneficial) action in question" (Jones, 1991, p. 376). This is in line with Hunt and Vitell's (1986) concept of "importance of stakeholders" as an influential variable on the ethical judgment of marketers. Jones posits that people care more about others who are close to them than they do for others who are distant. The positive impact of proximity on ethical recognition and behavior has been observed in several studies (Leitsch, 2006). Therefore, we posit that

H6: *In the context of a potential DA discriminatory recommendation, users who perceive more proximity toward the subjects of their decision are more likely to recognize the moral aspect of the issue.*

H7: *In the context of a potential DA discriminatory recommendation, users who perceive more proximity toward the subjects of their decision are less likely to approve the discriminatory recommendation.*

METHODOLOGY

This study will devise a single factor experimental approach by manipulating the level of demographic transparency between participants such that each participant is randomly assigned to one of the three groups: (i) DT-0: no demographic transparency; (ii) DT-1: demographic transparency in the form of charts; and (iii) DT-2: demographic transparency in the form of charts accompanied by representative images. A fictitious DA tool will be developed that will include more than 200 records of individuals who seek to buy a house. The system will then recommend a group of about 50 individuals to receive the promotion for a specific house. This sample will include potential discrimination against female buyers similar to the one depicted in Figure 2.

It is noteworthy that representative images to be used in this study are not images of the subjects in the dataset but are sample images that represent the demographic class well and can be bought and/or taken from volunteer individuals. Also, this study uses a collage of individuals as opposed to a photo of one single individual in order to minimize the impact of other variables that might contaminate the results (e.g., misunderstanding the single images in the context of female vs. male comparison to be related to race). Such an approach will also minimize the possible impact of participants' perceived homophily (toward an image of a person with the same race, etc.).

To ensure content validity, all measurement instruments will be adapted from existing and validated scales. Recognition of the moral issue, perceived proximity, mindfulness, and ethical culture will be measured using Reynolds' (2006) 3-item, Barnett's (2001) 3-item, Brown Ryan's (2003) 15-item, and Treviño et al. (1998) 14-item scales respectively. In addition, this study will control for the effect of participants' gender, age, social desirability responses, and knowledge of the housing industry as well as their prior knowledge of the possibility of DA tools generating discriminatory recommendations.

The sample for this study will consist of middle managers. According to a power analysis, 108 subjects would assure a sufficient statistical power of 0.80 to detect a medium effect size (f= 0.25). However, since this study will use PLS for data analysis, a minimum required sample size will be 150 (ten times the number of items of the construct with the highest number of items) (Gefen, Straub and Boudreau, 2000). Partial Least Squares and analysis of variance (ANOVA) techniques will be employed to test the research model and hypotheses.

CONCLUSION

The main goal in this research is to devise a method to alleviate the problem of discriminatory decision making when using DA tools. To that end, this study conceptualized and operationalized the notion of demographic transparency as a means of providing DA users with aggregated demographic information about the subjects of their DA-aided decisions. This study stands to make significant contributions to research. First, to the best of our knowledge, this is the first empirical study that examines the issue of discriminatory decision making using DA tool and strives to ameliorate the problem by focusing on the human aspects of decision making in using such tools. Previous studies have either produced generalized suggestions that using those tools can lead to making discriminatory decisions or merely focused on the technical antecedents of such discriminatory recommendations. In addition, the present study integrates three streams of research, data analytics literature, business ethics literature, and DSS guidance studies to be the first empirical study that will reveal the impact of availability and design of aggregated demographic information on DA users' perceptions and behavior.

This study also has significant potential practical implications. The introduction and operationalization of the two levels of DT can help organizations deal with the issue of unintentional discriminatory decisions made by DA users. In addition, the important role of users' mindfulness can help practitioners who seek to lower the level of discriminatory decisions made by their DA users as mindfulness is not a stable trait and can be increased by training programs (Kabat-Zinn, 1982). Organizations can also cultivate an ethical culture that encourages employees to engage in ethical decision making.

Several limitations also exist for this study. First, participants for this study will only be recruited in North America. Second, this study only focuses on discrimination against one demographic category (i.e., gender). Finally, this study will be conducted in a context in which participants use DT for the first time. Therefore, further research is warranted to test the impact of DT on users when they are repeatedly exposed to it.

REFERENCES

1. Amit, E. and Algom, D. and Trope, Y. (2009). Distance-dependent processing of pictures and words. *Journal of Experimental Psychology: General, 138*(3), 400-415.
2. Bar-Anan, Y. and Liberman, N. and Trope, Y. (2006). The association between psychological distance and construal level: evidence from an implicit association test. *Journal of Experimental Psychology: General, 135*(4), 609-622.
3. Barnett, T. (2001). Dimensions of moral intensity and ethical decision making: An empirical study. *Journal of Applied Social Psychology, 31*(5), 1038-1057.
4. Brown, K. W. and Ryan, R. M. (2003). The benefits of being present: mindfulness and its role in psychological well-being. *Journal of personality and social psychology, 84*(4), 822-848.
5. Craft, J. L. (2013). A review of the empirical ethical decision-making literature: 2004–2011. *Journal of Business Ethics, 117*(2), 221-259.
6. Crawford, K. (2013). The hidden biases in big data. *Harvard business review*. Retrieved from https://hbr.org/2013/04/the-hidden-biases-in-big-data
7. Ebrahimi, S. and Ghasemaghaei, M. and Hassanein, K. (2016). *Understanding the Role of Data Analytics in Driving Discriminatory Managerial Decisions*. Paper presented at the 37th International Conference on Information Systems, Dublin, Ireland.
8. Feldman, M. and Friedler, S. A. and Moeller, J. and Scheidegger, C. and Venkatasubramanian, S. (2015). *Certifying and removing disparate impact*. Paper presented at the Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Sydney, NSW, Australia.
9. Gefen, D. and Straub, D. and Boudreau, M.-C. (2000). Structural equation modeling and regression: Guidelines for research practice. *Communications of the Association for Information Systems, 4*(7), 1-70.
10. Hong, Y.-y. and Morris, M. W. and Chiu, C.-y. and Benet-Martinez, V. (2000). Multicultural minds: A dynamic constructivist approach to culture and cognition. *American psychologist, 55*(7), 709-720.
11. Hunt, S. D. and Vitell, S. (1986). A general theory of marketing ethics. *Journal of macromarketing, 6*(1), 5-16.
12. Jones, T. M. (1991). Ethical decision making by individuals in organizations: An issue-contingent model. *Academy of management review, 16*(2), 366-395.
13. Kabat-Zinn, J. (1982). An outpatient program in behavioral medicine for chronic pain patients based on the practice of mindfulness meditation: Theoretical considerations and preliminary results. *General hospital psychiatry, 4*(1), 33-47.
14. Leitsch, D. L. (2006). Using dimensions of moral intensity to predict ethical decision-making in accounting. *Accounting Education: an international journal, 15*(2), 135-149.
15. Newell, S. and Marabelli, M. (2015). Strategic opportunities (and challenges) of algorithmic decision-making: A call for action on the long-term societal effects of 'datification'. *The Journal of Strategic Information Systems, 24*(1), 3-14.
16. Pedreshi, D. and Ruggieri, S. and Turini, F. (2008). *Discrimination-aware data mining*. Paper presented at the Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining.
17. Podesta, J. and Pritzker, P. and Moniz, E. J. and Holdren, J. and Zients, Z. (2014). *Big data: Seizing opportunities, preserving values*. Retrieved from Washington, DC:
18. Rest, J. R. (1986). *Moral development: Advances in research and theory*. New York: Praeger.
19. Reynolds, S. J. (2006). Moral awareness and ethical predispositions: investigating the role of individual differences in the recognition of moral issues. *Journal of applied psychology, 91*(1), 233-243.
20. Ruedy, N. and Schweitzer, M. (2010). In the Moment: The Effect of Mindfulness on Ethical Decision Making. *Journal of Business Ethics, 95*(1), 73-87.
21. Silver, M. S. (1991). Decisional guidance for computer-based decision support. *MIS Quarterly, 15*(1), 105-122.
22. Singhapakdi, A. and Vitell, S. J. and Kraft, K. L. (1996). Moral intensity and ethical decision-making of marketing professionals. *Journal of Business Research, 36*(3), 245-255.
23. Street, M. D. and Douglas, S. C. and Geiger, S. W. and Martinko, M. J. (2001). The Impact of Cognitive Expenditure on the Ethical Decision-Making Process: The Cognitive Elaboration Model. *Organizational behavior and human decision processes, 86*(2), 256-277.
24. Treviño, L. K. and Butterfield, K. D. and McCabe, D. L. (1998). The ethical context in organizations: Influences on employee attitudes and behaviors. *Business Ethics Quarterly, 8*(3), 447-476.

25. Trope, Y. and Liberman, N. (2010). Construal-level theory of psychological distance. *Psychological review, 117*(2), 440-463.