Quantifying Biases in Social Media Analysis of Recreation in Urban Parks

Afra Mashhadi
Computer Systems and Software Division
University of Washington
Bothell, WA, USA
mashhadi@uw.edu

Samantha G. Winder
EarthLab
University of Washington
Seattle, WA, USA
sgwinder@uw.edu

Emilia H. Lia
EarthLab
University of Washington
Seattle, WA, USA
emmilia@uw.edu

Spencer A. Wood
EarthLab, eScience Institute
University of Washington
Seattle, WA, USA
spwood@uw.edu

***Abstract*—Recent years have seen an increase in the use of social media data for various decision-making purposes in the context of Urban Computing and smart cities, including management of public parks. Parks are critical natural assets that promote the health and well-being of urban residents, making it important that managers consider information about the quantity and character of park use in location and design decisions. However, as policy and management decisions rely on more autonomous methods, a critical concern that arises is the extent to which such analysis is fair and inclusive. In this article, we examine the biases that exist in data that are commonly used for the purpose of quantifying recreational use of urban parks. More precisely, we demonstrate the biases that exist in different sources of social media by comparing posts that are shared on Instagram and Flickr from ten urban parks in Seattle, WA. We compare the extent to which these platforms differ in terms of the information they capture about the number of people that visit the selected locations. We then demonstrate how further biases may be imposed when leveraging artificial intelligence to detect the count and demographics of park visitors, by comparing against an intercept survey of visitors.**

I. Introduction

In recent years Urban Computing [32], [28] has leveraged the proliferation of location-based social media to offer a fresh perspective of how cities function. Recent research showing that social media and other user-generated data sources provide reliable information about people in urban environments has caught the attention of practitioners and policy-makers who aim to make cities more livable and equitable. Practitioners are especially attracted to social media and other user generated content (UGC) because they are touted as a cheap and instant source of data on people in urban environments. However, there is an (understudied) risk that underlying biases in how these data are generated or analysed could lead decision-makers to unknowingly implement inequitable policies.

Location-based social media offer myriad insight into the social dynamics of cities and interactions that people have with urban environments [6], [22]. Additionally, these data often include valuable opinions that can be used to study the well-being of urban societies, as performed by De Choudhury et al. [7]. Social media data can be explored to study the social and economic characteristics of city dwellers [30]. Similarly, as we have shown in our past research, the images posted on social media platforms and their geo-tagged locations are a reliable source of data on the home locations of visitors to parks and other recreation destinations [31], offering an understanding of the socio-economic background of these visitors. As these types of analyses are becoming more popular and more cost effective than traditional surveys, policy makers are moving towards more autonomous human-centric sensing that includes, but is not limited to, social media analyses. However concerns have also started to arise regarding how accurately the social media data and analyses portrays the reality of our cities.

If we are to use social media data for planning and management that promotes sustainability and equity of urban cities, we first need to know who produces this data and what portion of the population is left out by it. Different social media platforms are known to attract unique demographic groups of users. For instance, research has shown that the Instagram and Pinterest user-base has more young females than males. Instagram appeals more to urban, African-American and Latina users, whereas Twitter accounts tend to belong to young, male and urban residents. The majority of Flickr users are male with a median age of 39[25]. Using multiple sources of big data has been proposed as a way to help overcome these biases [29], [9], but with little concrete evidence that it can work in practice.

In addition to data and population biases, artificial intelligence (AI) algorithms that are used in order to extract features from datasets could also introduce biases that could further skew the results [12], [23], [21]. For example, Bolukbasi et al. has shown that the popular word embedding space, Word2Vec, encodes societal gender biases [2], and some face recognition algorithms have been shown to misidentify people of color and women at high rates [17].

In this paper we study the viability of analyzing social media data as a means to capture the number and demographics of visitors to city parks in Seattle, WA. Urban parks offer unique opportunities for recreation and leisure activities which are vital to physical, social, and psychological health [11], [16]. In order to maintain the recreational benefits of parks, managers rely on information on the amount and character of park use to answer complicated questions about what services to provide, how to prioritize maintenance, and where to site new parks. In this study, we explore the potential for geo-tagged images that are publicly shared on two social media platforms to estimate the number and demographic characteristics of visitors. We ask whether geo-located images shared publicly on social media can accurately portray visitation and visitor demographics. Alternatively, are the conclusions drawn from social media biased in a way that would lead practitioners to unwittingly perpetuate existing inequalities in the opportunities and benefits of outdoor recreation to urban residents?

II. Background

A. Recreational Studies

The idea that crowd-sourced and ubiquitous data from social media can provide insights into patterns of park visitation, as well as the characteristics and behaviors of park visitors, has been the topic of recent research [13]. Studies have concluded that counts of park visitors are correlated with the popularity of the same destinations on social media platforms such as Flickr, Twitter, and Instagram [8], [27], [29], [10]. Social media popularity is typically quantified in photo-user-days (PUD), or the number of unique visitors who post a photograph online from a particular park on a given day [31]. Related studies have observed that the home locations of park visitors can also be inferred from social media, based on the information shared in users' public profiles [31], [27], or the locations of other content that each user shares publicly on the platform [20]. In this way, the locations of photographs that are shared online can be a reliable source of data on travel routes of park visitors across multiple destinations [15]. Additionally, home location data may be used to infer visitor demographics in order to understand where parks meet the demand for public spaces in ways that are fair and equitable. A limited number of studies have analysed the content of social media in order to map spatial patterns in recreational activities such as fishing that are apparent in images [26], [18]. To the best of our knowledge, however, there are no studies that have estimated visitation rates or visitor demographics from image content.

B. Demographics Studies Using Social Media

In order to estimate the demographics of the visitors we require a technique to automatically detect gender, age and ethnicity. However, due to privacy considerations, social-media accounts are often devoid of these basic demographic data. Techniques such as user profiling and demographic collators offer one way to retrieve these data by combining information from multiple sites and accounts for a given user, but these may also lead to ethical concerns.

Alternatively, and according to [3], there are currently two primary techniques used to infer demographics: one is to process the textual content of a user's profile such as their name and profile description to detect their demographics. However this method is known to be error-prone due to nicknames or arbitrary usernames, and the language used in providing the description of themselves is likely to follow a formal template, making it hard to detect age and the language skill of the writer. Furthermore, such techniques limit results to the demographic information of those individuals who are the members of the social media platforms. This excludes individuals who might appear in the photographs but are not members of the studied social media platform — such as children, elderly, or other specific populations. The second technique is to process the content of the images by leveraging progress in image recognition. In this vein, Jung et al. [14] has shown that the Face++ algorithm has an accuracy of 0.93 in detecting the ethnicity of 100 celebrities from IMDB. However, the authors also highlight that this high accuracy might be the result of the high-quality image dataset. We employ the latter technique and use the Face++ API to detect age, gender and race of those who appear in a photographs posted to social media.

III. Methods

This study investigates the prevalence of two types of biases that arise in urban computing analyses that rely on social media data: i) those that originate from how the data are generated on a social media platform and ii) those that are caused by the algorithms used to analyze these data. In order to understand these two types of biases, we employ several data sources in a comparative study that focuses on two metrics that are commonly used by recreation planners as the basis for management decisions policies: visitation rate and visitor demographic information. The study examines data from ten city parks in Seattle, WA.

A. Site Selection

Ten focal parks were randomly selected to be representative of a broad range of park types, neighborhoods, and user-groups across the city of Seattle. To ensure that we included parks from a variety of demographic areas, we stratified using the social vulnerability index (SVI) developed by the Centers for Disease Control and Prevention (CDC). This index combines 15 US census variables grouped into four themes (socioeconomic status, household composition, race/ethnicity/language, and housing/transportation) in order to rank census tracts by their relative vulnerability to hazardous events. We classified each park managed by the City of Seattle from 1 (low SVI) to 5 (high SVI) according to its location and randomly selected 2 parks from each category for inclusion. In order to confirm that study parks would have sufficient social media for our analysis, we dropped any parks with fewer than three average annual posts to Flickr.

B. Data Sources

In this study we rely on two sources of data from social media and one on-site visitor intercept survey. We use the number and content of images shared from Seattle City Parks on Instagram and Flickr as data for estimating visitation rates and demographics of visitors.

We used the Instagram *graphql* API to collect every image that was shared publicly and assigned to an Instagram locations within one of the selected Seattle parks (Table I). Since Instagram no longer provides an API endpoint for querying

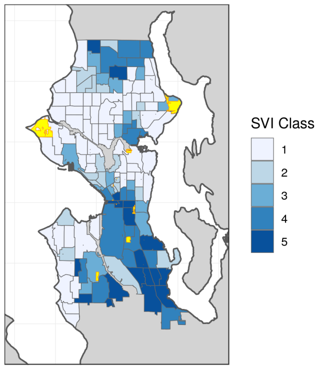

Fig. 1. Social Vulnerability Index classes by census tract in Seattle (WA). The yellow regions depict the selected parks.

locations, we first manually searched for locations in the Instagram web interface, using the park name and major features as search terms. The number of Instagram locations that were available for any given park were correlated with park area, presumably because bigger parks have more locations for users to tag. We collected images that were uploaded from January 2016 to January 2019. Metadata available for each image include the Instagram users' identifier and the date that the image was shared.

Park Name	Instagram	Flickr
City Hall	168	133
Discovery	41994	7154
Hing Hay	2030	501
Jefferson	2737	559
Judkins	2453	164
W. Magnuson	44235	15771
Montlake	1022	152
Plymouth Pillars	1007	117
Riverview	341	119
Summit	364	250

TABLE I
THE NUMBER OF COLLECTED INSTAGRAM AND FLICKR IMAGES PER STUDY SITE FROM JANUARY 2016–JANUARY 2019.

We queried the Flickr API for all geo-located photographs that were taken within the bounds of the ten study parks from January 2016–September 2019 (Table I). These photographs contain metadata including a unique user identifier, date the photograph was taken, and the latitude/longitude location of the camera when the image was taken. The geo-location typically comes from a GPS in the camera, but may also be manually assigned by the user by zooming and clicking on a webmap as the user uploads photographs to Flickr.

Between April–July, 2019, we conducted exit interviews at the study sites. Every park was surveyed on two weekdays and one weekend day, once in the morning, afternoon, and evening, for approximately 4 hours each. During this time, the interviewer intercepted visitors at five randomly selected exits, and asked them to complete a written survey in English. Participation was voluntary and no compensation was provided. The survey contained questions about visitors' activities, demographics, experiences, and feelings about parks in their neighborhoods. Our detailed survey questionnaire can be found in [19]. We collected 165 surveys in total. The number of responses ranged from seven to 39 surveys per park. Across all survey periods, the overall response rate was 16% of exiting visitors, although surveyors were unable to approach every exiting visitor. Approximately 41% of park users who were approached agreed to take the survey.

C. Metrics

We measure visitation by computing PUD per park, as it is the most widely used metric in the urban recreation literature [8], [31]. PUD is a metric that is used to capture the unique number of users posting photographs in a specific location per day. We compare PUD across the two data sources (Instagram and Flickr) to demonstrate the data biases associated with each platform. As we have no means of collecting self-identified demographic information from Flickr or Instagram users (e.g., by directly contacting them), we follow the methodology described in [24] and bound our definition of demographics to binary values of *white (vs non-white)* and *children (vs adults)*. We chose to do so because racial and ethnic identities are complex and evaluations by others may not match an individuals' self-identification, so asking crowd-sourced workers to classify race or ethnicity of individuals in photographs (beyond the binary values defined above) did not feel appropriate. We assume that a person is a child if their estimated age by Face++ is less than 16 years old.

To serve as labelled data, we used Amazon Mechanical Turk to ask workers to perform a simple counting task in the context of image analysis. Specifically, we asked them to count the total number of people, the number of white people, and the number of children that appeared in each photo. In hiring the crowd workers we followed minimum wage regulation of our state (14 USD per hour). Our labelled dataset is composed of 500 images randomly selected across the ten parks from both Instagram and Flickr.

We then relied on the labelled data to measure the algorithmic biases. We compute the recall and precision of the algorithm in terms of the accuracy of its visitor detection. We used the labelled dataset and define precision as the fraction of photographs with an accurate count of visitors appearing the image (i.e., matching the label) among all the photographs that the algorithm retrieved (i.e., counted at least one person in them). We define recall as the fraction of the photographs with an accurate count among all images (including images in which the algorithm did not detect any people).

IV. RESULTS

A. Platform Biases

Comparing the two sources of social media, we find that PUD – a metric commonly used to study recreational visits – is substantially higher according to Instagram than Flickr. Park visitors in Seattle are more likely to share content about their experience on Instagram compared to Flickr. Figure 2 demonstrates this result in terms of weekly PUD across all ten study locations. Furthermore, for both platforms the long-term trend in PUD appears to reflect the popularity of each platform, with Flickr following a downward trend and Instagram becoming more popular over time.

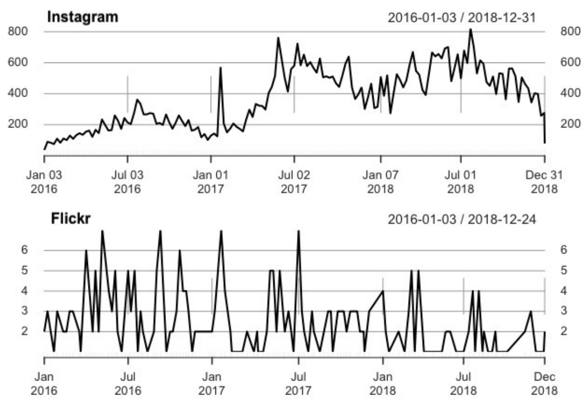

Fig. 2. Weekly total number of users per day who posted photographs to Instagram (top) and Flickr (bottom) from the ten studied parks from January 2016–January 2019.

Responses to our on-site survey support the observation that social media platforms differ in their popularity, where 59% of survey respondents stated that they use social media. Of these, 62% said that they share content on Instagram, as opposed to only 2% who said they share content on Flickr. When asked what they might share to social media from their trip today, the most popular response was a "scenic view", followed by "pets" and "wildlife" (Figure 3). Eighty percent of those who use social media said that they might post content with information about other people (either friends, family, sports, or other event or gathering). Respondents were allowed to select multiple topics.

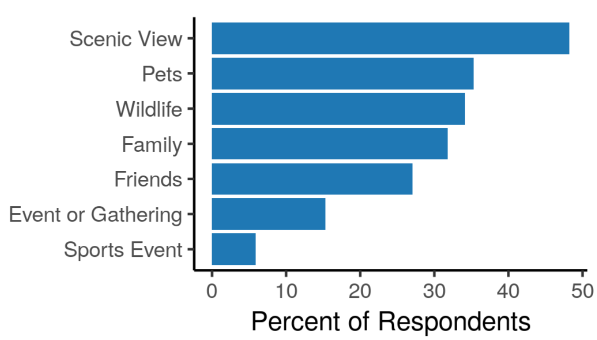

Fig. 3. Topics of photographs that respondents who use social media report that they might share during visits to urban parks.

Turning to the content of the images posted to Instagram and Flickr, we observe that the average number of people that appear in photographs is consistent across sites and between social media platforms (Figure 4, presenting the mean and standard deviation for samples of 50 photographs from four most visited parks). A t-test also indicates that there are no significant differences in the average number of people that appear in images. The same human-labelled dataset also shows that 39% of the people appearing in the images are children (mean= 0.28 per photo, sd=0.40). This result differs from the proportion of children as reported in our on-site survey, where respondents reported the number of adults and the number of children in their party. Across our 165 respondents, the total number of adults was 250, and the total number of children was 46: a lower ratio than what we observe from our images.

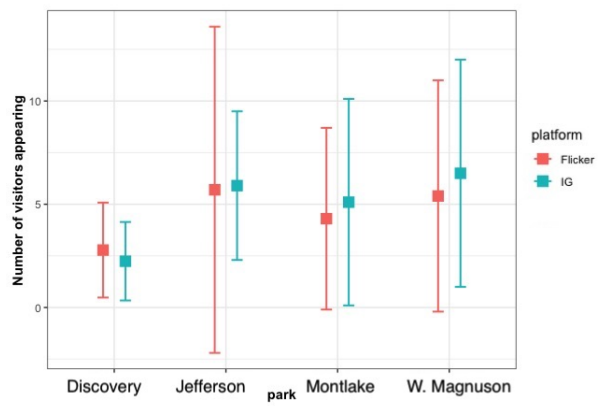

Fig. 4. Average number of people who appear in 50 photographs randomly selected from each of the four most popular city parks based on the labelled data. Bars show standard errors.

In our comparison of the demographics of park users, we find differences in the ages of visitors detected by Face++ in photographs posted to Instagram as opposed to Flickr (Figure 6). There is slight downward shift in the age distribution of people in Flickr images, suggesting that Flickr images contain a higher number of children. The racial composition of people in the same images, according to Face++, is also different across the two platforms ($\tilde{\chi}^2 = 14.25$ (p= 0.002)). Similarly, the racial composition of the visitors who appear in Instagram photographs from 2019 also differs from the groups that were reported by survey respondents (Figure 5 illustrates this comparison). A chi-squared test examining the relationship between race and detection method (Face++ vs Survey) found that the results varied by detection method ($\tilde{\chi}^2 = 8.9$, p = .01), using only 2019 Instagram photographs in order to overlap with the time period of our on-site intercept survey. A greater proportion of survey respondents were white than identified by Face++.

B. Algorithmic Biases

In order to evaluate the biases that are potentially caused by algorithmic methods for estimating the park visitation and visitor demographics, we first measure the performance of the Face++ algorithm based on precision and recall of the estimated number of people in images, compared to human-labelled data. Similar to previous works we find precision, the fraction of photographs where the number of visitors was

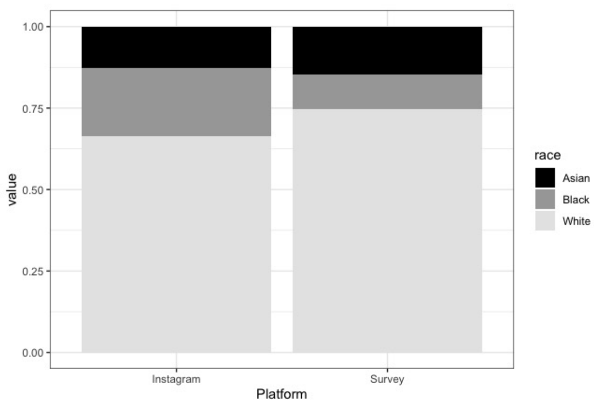

Fig. 5. Race of park visitors based on Face++ detection of people in Instagram images (A) and on-site survey respondents (B) in 2019.

correctly classified out of photographs with at least one person identified, to be high (0.91). However we find the recall, the proportion correctly classified out of all photographs, is only 0.52. This indicates that many visitors were undetected and that Face++ would underestimate park visitation rates. The low recall also provides evidence that the algorithm is significantly impacted when it is applied to datasets in-the-wild instead of the celebrity datasets that are used to train it.

Face++ continuously under-counts the number of people who appear in photographs (Figure 7). The number of people that are counted in each photograph by Face++ is consistently lower than the count given by crowd who view the same image. In order to understand how specific demographic characteristics (age and race) contribute to these algorithmic biases we bin each image according to whether the algorithm either under-counted, over-counted, or correctly determines the number of people in the image. The algorithm correctly estimates the number of people who are present in 51% of the photographs. In 45% of the images, the algorithm detects fewer people than were actually present. The algorithm over-detects the number of people in 4% of the images. Examining this small subset of pictures we find that almost all the cases corresponded to pets' faces. Since 35% of our survey respondents said that they might share images of their pets on social media, this has the potential to be a substantial source of bias in some locations.

To investigate which demographic features of the visitors is contributing to the detection error of the algorithm, we ask at whether race (in our case categorized in two broad categories of white vs non-white) contributes to the number of visitors not being detected by Face++. Assuming the null hypothesis that the ratio of white people appearing in the photographs is independent of the detection rate, we compare the ratio of white people that appear in photographs in the under-detected bin with the ratio of white people that appear in the photographs from the accurately-detected bin. We do not find statistically significant evidence to reject the null hypothesis. That is we cannot say that the under-detection bias is related to perceived race of the subjects. However, this result does not imply that Face++ is accurate in detecting the race of the visitors. Based on the accurately detected bin, a test of equal proportions reports that Face++ identifies a significantly higher percentage of white subjects per photograph (mean=0.64) than the human-labelled data created by Mechanical Turk workers (mean=0.31). That is, the algorithm incorrectly classifies some non-white subjects as being white. Based on our visual observation we speculate that this is due to the nature of photographs in the parks, where subjects' faces could be obscured by sunglasses or sporting accessories, for instance.

The relationship between under-detection and the age of the visitors indicates that there is an algorithmic bias created by the failure to recognize children in images. A test of equal proportions shows significant differences in the mean percentage of kids appearing in under-detected (mean=0.31) photographs versus the accurately detected photographs (mean=0.13). Looking at the under detected bin, we find a correlation between the ratio of the children present in the images and the detection error of correlation = 0.37, $p < 0.001$).

V. Discussion

In this paper we present the results of multiple approaches that rely on Instagram and Flickr photographs to estimate the number of visitors to Seattle urban parks. We showed that the popularity of the platforms can heavily skew the visitation count. It is important to note that Instagram and Flickr are not unique cases, indeed as researchers and policy makers are heavily investing their efforts to use social media analyses caution needs to be taken on how data source biases are handled. We have also presented that the advances in face recognition techniques could be leveraged to enable recreational policy makers to remove some of the social media platform biases (i.e, popularity) and demographic biases of the intercept surveys.

However our analysis demonstrated that such techniques must be used with extreme caution due to the algorithmic biases that could interfere with the results. In this vein, we showed that the AI consistently under-counts the number of visitors, and that this bias mostly impacts children. We also presented that while the detection error is not directly correlated with race, nonetheless the algorithm suffers from racial bias by consistently classifying non-white individuals as white.

A. Limitation

We acknowledge that our work has the following limitations. Firstly the AI algorithm we used served as black box and we do not know on what data it was trained. Secondly the AI algorithm did not classify races such as Hispanic and Native American, which in some areas of the city we studied are majority minority ethnicity. Although in our survey we collected this information we did not report these results due to the mismatch with the Face++ categories. This limitation also

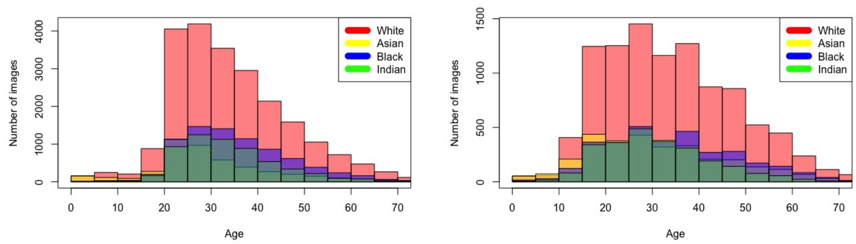

Fig. 6. Age of people detected by Face++ algorithm and for various ethnicity for Instagram and Flicker. respectively.

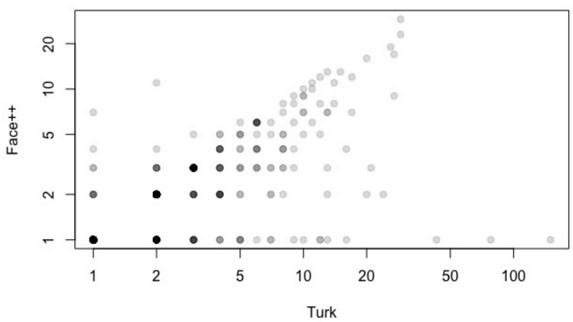

Fig. 7. Number of visitors detected in images according to human observers and the Face++ algorithm.

resonates with other works in computational social science were an existing challenge is the definition and boundary of ethnicity. For example in [4], [5] authors both used the term ethnicity to refer to a classification system that includes both racial and ethnic identities (black, white, asian, hispanic) where as others [1] used the same classification system but addressed it as *race*.

B. Implications

Our study has the following practical and theoretical implications. From practical standpoint, our study highlights the biases that can surface when social media analysis is used for recreational and policy making purposes. One of the main findings of this study was the striking similarity in terms of crowd count for both platforms, suggesting the need to migrate from traditional visit counts metrics to a more platform agnostic one. Our result also highlighted how the most vulnerable group of people, in this case kids, are those that are excluded due to algorithmic biases.

From the theoretical perspective, first and foremost our work highlights the need for future research direction on domain adaptation and knowledge transfer. Traditional machine-learning algorithms such as the supervised learning one that we used in this study train statistical models to make predictions on unseen data. However, these models do not guarantee optimal performance if the test data are vastly different from the training data. A main approach that is currently under investigation in Machine Learning community is to pay more attention to reducing the effort involved in recollecting labeled data and retraining a new model by using knowledge transfer between domains. However, as we demonstrated in this work collecting in-the-wild demographics labels is extremely sensitive topic which requires further investigation.

Acknowledgments

We would like to thank Todd Burley who provided insights and information on Seattle parks, Monika Derrien and Dale Blahna who helped with the desgin of the intercept survey and their valuable feedback on this project. The funding for this project was supported by the Bullitt Foundation.

References

[1] E. M. Ardehaly and A. Culotta, "Using county demographics to infer attributes of twitter users," in *Proceedings of the joint workshop on social dynamics and personal attributes in social media*, 2014, pp. 7–16.

[2] T. Bolukbasi, K.-W. Chang, J. Y. Zou, V. Saligrama, and A. T. Kalai, "Man is to computer programmer as woman is to homemaker? debiasing word embeddings," in *Advances in neural information processing systems*, 2016, pp. 4349–4357.

[3] N. Cesare, C. Grant, and E. O. Nsoesie, "Detection of user demographics on social media: A review of methods and recommendations for best practices," *arXiv preprint arXiv:1702.01807*, 2017.

[4] J. Chang, I. Rosenn, L. Backstrom, and C. Marlow, "epluribus: Ethnicity on social networks," in *Fourth International AAAI Conference on Weblogs and Social Media*, 2010.

[5] X. Chen, Y. Wang, E. Agichtein, and F. Wang, "A comparative study of demographic attribute inference in twitter," in *Ninth International AAAI Conference on Web and Social Media*, 2015.

[6] J. Cranshaw, R. Schwartz, J. Hong, and N. Sadeh, "The livehoods project: Utilizing social media to understand the dynamics of a city," in *Sixth International AAAI Conference on Weblogs and Social Media*, 2012.

[7] M. De Choudhury, S. Sharma, and E. Kiciman, "Characterizing dietary choices, nutrition, and language in food deserts via social media," in *Proceedings of the 19th acm conference on computer-supported cooperative work & social computing*. ACM, 2016, pp. 1157–1170.

[8] M. L. Donahue, B. L. Keeler, S. A. Wood, D. M. Fisher, Z. A. Hamstead, and T. McPhearson, "Using social media to understand drivers of urban park visitation in the twin cities, mn," *Landscape and Urban Planning*, vol. 175, pp. 1–10, 2018.

[9] D. M. Fisher, S. A. Wood, Y.-H. Roh, and C.-K. Kim, "The geographic spread and preferences of tourists revealed by user-generated information on jeju island, south korea," *Land*, vol. 8, no. 5, p. 73, Apr 2019.

[10] D. M. Fisher, S. A. Wood, E. M. White, D. J. Blahna, S. Lange, A. Weinberg, M. Tomco, and E. Lia, "Recreational use in dispersed public lands measured using social media data and on-site counts," *Journal of Environmental Management*, vol. 222, pp. 465–474, Sep. 2018.

[11] H. Frumkin, G. N. Bratman, S. J. Breslow, B. Cochran, P. H. Kahn, J. J. Lawler, P. S. Levin, P. S. Tandon, U. Varanasi, K. L. Wolf, and S. A. Wood, "Nature contact and human health: a research agenda," *Environmental Health Perspectives*, vol. 125, p. 075001, 2017.

[12] D. J. Fuchs, "The dangers of human-like bias in machine-learning algorithms," *Missouri S&T's Peer to Peer*, vol. 2, no. 1, p. 1, 2018.

[13] A. Ghermandi and M. Sinclair, "Passive crowdsourcing of social media in environmental research: A systematic map," *Global Environmental Change*, vol. 55, pp. 36–47, Mar. 2019.

[14] S.-G. Jung, J. An, H. Kwak, J. Salminen, and B. J. Jansen, "Assessing the accuracy of four popular face recognition tools for inferring gender, age, and race," in *Twelfth International AAAI Conference on Web and Social Media*, 2018.

[15] B. L. Keeler, S. A. Wood, S. Polasky, C. Kling, C. T. Filstrup, and J. A. Downing, "Recreational demand for clean water: evidence from geotagged photographs by visitors to lakes," *Frontiers in Ecology and the Environment*, vol. 13, no. 2, pp. 76–81, Mar. 2015.

[16] Keeler Bonnie L., Hamel Perrine, McPhearson Timon, Hamann Maike H., Donahue Marie L., Meza Prado Kelly A., Arkema Katie K., Bratman Gregory N., Brauman Kate A., Finlay Jacques C., Guerry Anne D., Hobbie Sarah E., Johnson Justin A., MacDonald Graham K., McDonald Robert I., Neverisky Nick, and Wood Spencer A., "Social-ecological and technological factors moderate the value of urban nature," *Nature Sustainability*, vol. 2, no. 1, p. 29–38, 2019.

[17] B. F. Klare, M. J. Burge, J. C. Klontz, R. W. V. Bruegge, and A. K. Jain, "Face recognition performance: Role of demographic information," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 6, pp. 1789–1801, 2012.

[18] H. Lee, B. Seo, T. Koellner, and S. Lautenbach, "Mapping cultural ecosystem services 2.0 – Potential and shortcomings from unlabeled crowd sourced images," *Ecological Indicators*, vol. 96, pp. 505–515, Jan. 2019.

[19] E. Lia, "Seattle city park visitor survey." [Online]. Available: https://ajahanba.github.io/seattleparks_survey_v4.pdf

[20] M. J. Martinez-Harms, B. A. Bryan, S. A. Wood, D. M. Fisher, E. Law, J. R. Rhodes, C. Dobbs, D. Biggs, and K. A. Wilson, "Inequality in access to cultural ecosystem services from protected areas in the Chilean biodiversity hotspot," *Science of the Total Environment*, vol. 636, pp. 1128–1138, Sep. 2018.

[21] S. U. Noble, *Algorithms of oppression: How search engines reinforce racism*. nyu Press, 2018.

[22] A. Noulas, S. Scellato, C. Mascolo, and M. Pontil, "Exploiting semantic annotations for clustering geographic areas and users in location-based social networks," in *Fifth International AAAI Conference on Weblogs and Social Media*, 2011.

[23] C. O'neil, *Weapons of math destruction: How big data increases inequality and threatens democracy*. Broadway Books, 2016.

[24] M. Pennacchiotti and A.-M. Popescu, "Democrats, republicans and starbucks afficionados: user classification in twitter," in *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2011, pp. 430–438.

[25] D. Quercia, L. M. Aiello, and R. Schifanella, "Diversity of indoor activities and economic development of neighborhoods," *PloS one*, vol. 13, no. 6, p. e0198441, 2018.

[26] D. R. Richards and D. A. Friess, "A rapid indicator of cultural ecosystem service usage at a fine spatial scale: Content analysis of social media photographs," *Ecological Indicators*, vol. 53, pp. 187–195, Jun. 2015.

[27] C. Sessions, S. A. Wood, S. Rabotyagov, and D. M. Fisher, "Measuring recreational visitation at U.S. National Parks with crowd-sourced photographs," *Journal of Environmental Management*, vol. 183, pp. 703–711, Dec. 2016.

[28] T. H. Silva, A. C. Viana, F. Benevenuto, L. Villas, J. Salles, A. Loureiro, and D. Quercia, "Urban computing leveraging location-based social network data: a survey," *ACM Computing Surveys (CSUR)*, vol. 52, no. 1, p. 17, 2019.

[29] H. Tenkanen, E. D. Minin, V. Heikinheimo, A. Hausmann, M. Herbst, L. Kajala, and T. Toivonen, "Instagram, Flickr, or Twitter: Assessing the usability of social media data for visitor monitoring in protected areas," *Scientific Reports*, vol. 7, Dec. 2017.

[30] A. Venerandi, G. Quattrone, L. Capra, D. Quercia, and D. Saez-Trumper, "Measuring urban deprivation from user generated content," in *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*. ACM, 2015, pp. 254–264.

[31] S. A. Wood, A. D. Guerry, J. M. Silver, and M. Lacayo, "Using social media to quantify nature-based tourism and recreation," *Scientific Reports*, vol. 3, p. 2976, 2013.

[32] Y. Zheng, L. Capra, O. Wolfson, and H. Yang, "Urban computing: concepts, methodologies, and applications," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 5, no. 3, p. 38, 2014.