Algorithmic Privacy and Gender Bias Issues in Google Ad Settings

Nisha Shekhawat*
nisha.shekhawat.y16@lnmiit.ac.in
The LNM Institute of Information Technology
Jaipur, India

Aakanksha Chauhan
aakanksha.chauhan.y16@lnmiit.ac.in
The LNM Institute of Information Technology
Jaipur, India

Sakthi Balan Muthiah
sakthi.balan@lnmiit.ac.in
The LNM Institute of Information Technology
Jaipur, India

ABSTRACT

For more than three years, Google has been facilitating users with four gender options – "Male", "Female", "Rather Not Say" and "Custom". "Rather Not Say" is for users who do not prefer to disclose their gender identity and "Custom" is for users who do not identify themselves among the conventional gender labels (male or female). By this, it is evident that Google provides choice to its users to classify themselves among non-conventional gender groups. This work makes an attempt to assess choice, transparency and privacy in Google Ad Settings when the option "Rather Not Say" is selected as gender. It was observed that even though the gender was set as "Rather Not Say", a conventional gender was displayed as demographic in Ad Personalization page of Google Ad Settings. Therefore, even though it provides choice to the user, it is not an absolute choice as Google still classifies an individual into one of the two traditional categories. Our experiment infers that the websites might be categorized as Male or Female-oriented. Therefore, while trying to create a preference of websites for a particular user, the system often introduces bias towards a gender for a predefined interest demographic in Google Ad Personalization page. This paper focuses on the statistical analysis of the prediction of gender for the different categories of websites and how this effects a user's choice, privacy and transparency.

CCS CONCEPTS

• **Security and privacy** → *Social aspects of security and privacy.*

KEYWORDS

Google ads, gender, bias, preference, transparency, privacy, choice and online browsing behaviour

ACM Reference Format:
Nisha Shekhawat, Aakanksha Chauhan, and Sakthi Balan Muthiah. 2019. Algorithmic Privacy and Gender Bias Issues in Google Ad Settings. In *11th ACM Conference on Web Science (WebSci '19), June 30-July 3, 2019, Boston, MA, USA.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3292522.3326033

WebSci '19, June 30-July 3, 2019, Boston, MA, USA

ACM ISBN 978-1-4503-6202-3/19/06…$15.00
https://doi.org/10.1145/3292522.3326033

1 INTRODUCTION

Google currently provides its users with four gender options – "Male", "Female", "Rather Not Say" and "Custom", which was first introduced in Dec 2014[2]. "Male" and "Female" options correspond to conventional gender demographics. The option "Rather Not Say" refers to the group of users who do not want to reveal their gender identity, and "Custom" is for the users who are part of the non-conventional gender groups. These policies are formed to provide users with a large domain of non-conventional options to opt from.

For every user, there are two categories of information that is collected by Google to facilitate the users with its services as stated in policies. First one is provided by the user while creating the account and second one is collected as the user employs the services provided by Google like web search, ads, geolocation, history, and many more as listed in Figure 1.

To cater transparency and choice, Google provides the service of Ad Personalization to the users. It displays the demographics which include, first, the information provided by the users, i.e., age group and gender, and second, the interests which are inferred by Google from the user's online browsing behaviour. This online browsing behaviour includes the website visits and web searches (keywords) conducted by a user. Google also provides the facility to disable these interest demographics if a user is not inclined towards them.

As the choice opted by a user is a major consideration by Google, this work attempts to develop an insight about the online browsing behaviour of the simulated users and how it is effecting their choice. This experiment shows that how the analysis of online browsing behaviour of a user, eventually, deprives him or her of the choice while making a selection for the sensitive attribute (gender), i.e., despite of choosing "Rather Not Say", a conventional gender is assigned. Furthermore, it also includes some observations that show how this gender prediction is thereby making the system less transparent for the user. Moreover, it can also be said that a user's privacy is hampered when this predicted information is shared with third party websites[5][7].

This paper is a study on two fronts: one it shows that the gender choice offered by Google is not absolute with respect to Google Ad scenario. And second, when Google assigns a gender in the Google Ad Settings it assigns with respect to the websites a Google user searches, visits or clicks – this shows that the algorithm follows the general societal perception of how male and female interests are stereo-typically differentiated. Both of these present various privacy and bias issues for the user. This work is to indicate that there is a bias and privacy issue associated with Google Ad Settings especially when the user selects "Rather Not Say" category of gender.

Prior work relevant to our study is discussed in Section 2. Method and categorization of outcomes obtained from this experiment are explained in Section 3. Section 4 emphasizes on the statistical analysis of outcomes obtained from this study. The issues detected in Google Ad Settings are discussed in Section 5 while conclusion and future work are explained in Section 6.

2 LITERATURE SURVEY

Several studies have been conducted on Google ads and the relevant works are discussed here. Datta et al.[3] examined Google ads to explore the effect of browsing behaviour on the displayed ads. They developed Adfisher tool to detect differences between the ads shown to the experimental and control groups. These groups had separately defined treatment for various simulated users. While Datta examined transparency for interest demographics in Ad Settings, our work majorly focuses on gender demographic. Craig E. Wills[10] assessed the ads shown to users during controlled browsing as well as examined the inferred demographics and interests shown in the Ad Preference Manager. They also observed profile based ads and found that Google does show non-contextual ads related to sensitive topics like health and finance.

The usage study closest to this work is that of Michael Carl Tschantz[8] who experimented to determine the accuracy of Google ad demographics. They monitored the user's browsing behaviour and recorded interests in Google Ad Settings to conclude whether inferred interests match with user-provided information. On the other hand, this paper studies patterns in user's online browsing behaviour which identifies the user as male or female and also concludes whether the algorithm is leading to a preference or bias. Small-scale anecdotal examinations of the accuracy of Ad Settings have appeared in the popular press, as has a survey looking at the accuracy of Google's geolocation abilities[8].

Other related works differ in both objectives and methodologies. They all focus on how visiting web pages change the ads seen. While our work examines which gender is assigned to a user, when the gender is set as "Rather Not Say", in accordance with the online browsing behaviour of the user and also makes an attempt to develop an insight into Google's algorithm which might be classifying websites as Male or Female-oriented.

3 METHODOLOGY

3.1 Google Ads Policy

As examined on January 2019[4], Google Ads Policy under the subsection of Data Transparency clearly distinguishes between "the information they collect as we use the services" and "the information we create or provide to them" as shown in Figure 1. The former includes the collection of data from browsing, videos, ads, location, website, and apps. This collected data serves as an alley to make the Google services[4] better for the users. An instance of the efficacy that Google provides is making Google Maps faster as the bits of a user's location are collected, which when combined with the location of other users, it is able to recognize a better traffic pattern to aid the users. The latter, whereas, is the information that a user provides while creating a Google account. Google also claims to protect this information as the user utilizes the services. This information includes name, date of birth (age), gender, phone number, and other sensitive data[4]. Privacy is one of the main concerns as stated in the policy and a comprehensive review process is promised, thereby making it a key consideration at each level of development of any service[4].

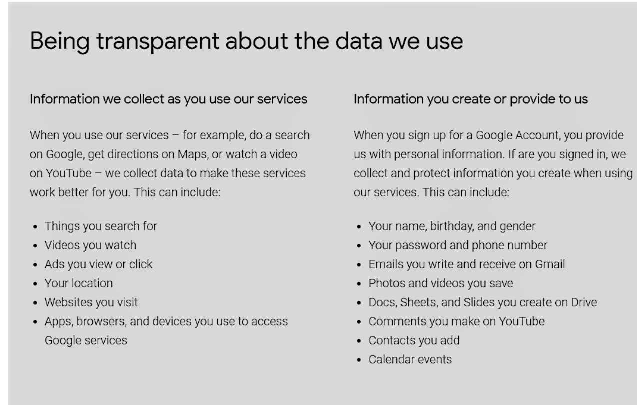

Figure 1: Google Ads Privacy Policy (https://safety.google/privacy/data/).

3.2 Approach

This experiment aims to assess the choice provided for gender demographic in Google Ad Settings by determining what kind of browsing behaviour pertains to a specific conventional gender according to Google ads algorithm. To conduct this experiment, fifty Gmail accounts were created manually. Gmail account creation includes mobile number verification and the number of accounts that can be validated with a unique mobile number is limited[6]. Therefore, automation of the Gmail account creation is not feasible.

This experiment is implemented using Firefox private browsing. Private browsing was necessary to ensure that no past history or bookmarks were stored that could effect observations of the current browsing behaviour of the simulated users. It was commenced by creating fifty Gmail accounts having same date of birth, no recovery email address or mobile number and gender being set as "Rather Not Say". At the beginning of the experiment, age group (18-24) was the only displayed demographic (inferred from date of birth) in the Ad Personalization page of Google Ad Settings. Each simulated user was then subjected to a different treatment and Ad Personalization page was monitored. The treatment for a simulated user involved searching a keyword, followed by visiting websites for that search. As Google ad profiles are dynamic and may change (e.g., new interests are added) therefore, Ad Personalization page is monitored after visiting each website. Outcomes for the above mentioned procedure are further stated in Section 4.1 and 4.2.

3.3 Categorization of collected information

The results that are collected by applying the approach (described in Section 3.2) to the Google accounts can be divided into two categories – "Label" and "Result". The first category holds record of the websites after each visit and states whether Ad Personalization page has displayed any changes in the interest demographics. This information is kept under the "Label" column which is classified

Table 1: Table for Label categorization

Username (@gmail.com)	Keywords	Treatment	Label
adfisher0002	web dev jobs	linkedin.com	Radical
		glassdoor.com	Radical
adfisher0003	highest paying finance jobs	cnbc.com	Steady
		businessinsider.com	Steady
		emolument.com	Radical

into "Radical" and "Steady". "Radical" is labelled when there is an addition of interest demographics in the Ad Personalization page, which is a consequence of the most recent and the preceding website visits of the simulated user. Whereas "Steady" is labelled when there is no addition or change in the Ad Personalization page after a web search and website visit of a unique user.

This categorization has been shown in Table 1[1] whose first column is Username of the Gmail account on which the treatment is done and second column is Keywords which are searched on web. The third column is Treatment which includes the websites that are visited from that account. The fourth column is Label which is the above mentioned categorization.

The other category is "Result" which is outcome of the complete treatment imparted to a unique account and is obtained from all demographics corresponding to that account. These are analyzed by the algorithm to respond with a gender which was supposed to be absent as "Rather Not Say" option was selected. On the basis of these outcomes, "Result" is classified into three groups i.e, Male, Female and No Gender. No Gender is stated if no result is acknowledged, even after subjecting an account to the complete treatment. For Male and Female classes, the treatment is completed when one of the two genders, either "Male" or "Female" is predicted. For No Gender, the treatment is supposed to be complete when an account visits a significant number (average of 7 to 10 sites) of websites to get a prediction but is still unable to classify one.

The categorization of Result is shown in Table 2[1] in which the first column is Username of the Gmail account, second column is Keywords of the web search, third column is Treatment that is provided to the simulated accounts, fourth column is Label that categorizes changes in interest demographics due to website visits into Radical and Steady. The fifth column is Result that shows the above stated categorization of the outcome of the treatment given to the accounts.
1

4 OBSERVATION

4.1 Observation of Label

"Label" holds the record of each web search and website visit, and whether the Ad Personalization page has responded with any change in the interest demographics or not. In this work, 86 out of 190 visited websites showed Radical behaviour, where 15 sites displayed Radical behaviour in the first visiting only whereas there were only 5 sites that showed Steady behaviour even after the treatment was completed. For instance, sites like *indeed.com, pic.nic.in* were labelled Radical because they displayed change in Ad Personalization page in their respective visits. On the other hand, sites

[1]https://github.com/nisha987/Google-gender

Table 2: Table for Result categorization

Username (@gmail.com)	Keywords	Treatment	Label	Result
adfisher0002	web dev jobs	linkedin.com	Radical	Female
		glassdoor.com	Radical	
adfisher0003	highest paying finance jobs	cnbc.com	Steady	Male
		businessinsider.com	Steady	
		emolument.com	Radical	
adfisher0015	parenting	parenting.com	Steady	No Gender
		psychologytoday.com	Radical	
		apa.org	Steady	
		ted.com	Radical	
		nytimes.com	Steady	

Table 3: Table of significant observations

S.No.	Keywords	Treatment	No. of simulated users	Result
1	yoga	yoga+online dating yoga+teaching +shopping	2	Female
	gym	gym gym+art of living	2	Male
2	buying apartments	apartment renting websites	2	Female
3	machine learning conferences	machine learning conferences	2	Male
		machine learning conferences+data science jobs	1	Female
4	data science jobs	data science jobs	4	Male
		machine learning conferences+data science jobs	1	Female
5	high paying finance jobs	finance jobs+chartered accountant	2	Male
6	software development	software development +data science	2	Male
		software developer+full stack developer	2	
		software developer +full stack developer	1	Female
		software developer+jewellery +soccer+home decor+cricket +nail art+bridal dress+pregnancy +pinterest	1	No Gender
7	disability parenting horoscope english literature alcohol support groups	disability+hospital parenting horoscope+demons english literature +political news alcohol support groups +stress management	8	No Gender
8	cooking and cooking jobs	chef jobs	1	Male
		cooking	1	Female
9	medical	doctor jobs	1	Male
		nurse jobs	1	Female

like *nhs.uk, parenting.com* were labelled Steady which means they did not immediately effect the Ad Personalization page.

4.2 Observation of Result

This experiment classified websites into three classes in accordance with general societal perception – Male-oriented, Female-oriented and Gender Neutral as shown in Table 4[1]. These categories were permuted to create a treatment comprising of websites from one or more of the above three classes. The inferred results of this experiment are stated in Table 3[1] in which the Keywords are being compared, along with the Treatment provided to various accounts, to show the number of simulated users that are showing deviation towards a single gender through their online browsing behaviour.

It was observed that if a user searches for career websites, the user was mostly predicted as "Male". Out of 21 searches for jobs, 13 were high paying professions, where 8 of these were predicted as "Male", 3 as "Female" and 2 accounts were not assigned any gender. For remaining 8 low paying professions, 7 were predicted as "Female" and 1 as "Male"[1]. It was startling to learn that visiting sites for cooking predicted that the user is female while searching for chef jobs predicted that the user is male. Two simulated users visited career websites of tech giants like Adobe, Google, Microsoft, IBM and Amazon searching for data science jobs and were classified as male. This implies that the skewed ratio of male to female employees is also effecting the job opportunities available to the female applicants. Another result demonstrated yoga as an interest for females and gym for males. Four simulated users visited expenditure associated websites and the algorithm concluded that the users were female. Bollywood and machine learning conferences represent male interests whereas flowers, home decor, embroidery, Bacardi, quit smoking, kitchen gardens, motivational talks serve as female interests in Google ad algorithms. For websites and web searches related to parenting, disability, alcohol support groups and horoscope, " No Gender" was concluded for Result column. We also observed that gender field along with other interests in Google Ad Personalization page changes with changing user browsing behaviour like inactivity (mostly vanishes) for a few weeks.

5 DISCUSSION

5.1 Choice

The absolute choice, when given to a user, refers to the fact that every option chosen by the user is respected by Google. In this experiment, for the gender "Rather Not Say", it is found that after a few website visits, a gender field appears in Ad Personalization page, identifying the user as Male, Female or No Gender on the basis of online browsing behaviour of the user (Section 3.3). The prediction of gender deprives the user from the facility of absolute choice.

5.2 Transparency

A user's online browsing behaviour is monitored and recorded to benefit him or her by improving the services as stated in the Google Ads Policy. But here, it can be analyzed that the property of transparency is not delivered entirely because this collected information, along with improving the services, also leads to prediction of gender (sensitive attribute), which according to the Google Ads Policy, is meant to be provided by the user. The prediction of gender, as observed in Section 4.2, might be a result of the classification of websites into either Male or Female-oriented or as Gender Neutral. This, eventually, might cater for bias while a user goes for any browsing activity. The classification that was observed from this experiment is stated in Table 4[1] .

Table 4: Categorization of websites

S. No.	Classes	Websites
1.	Female-oriented	Web Development Jobs – LinkedIn, Glassdoor Fashion designer jobs – LinkedIn, shine, indeed Nurse jobs – aetnacareers, glassdoor, indeed Online dating Yoga – verywellfit.com Hairdresser – Naukri, indeed, glassdoor Cooking Flowers Myntra, streetstyle Home decor, embroidery, nykaa Novels Adventure sports like skydiving, canoeing, desert camping, paragliding Kitchen gardens Teaching jobs and nursery rhymes Apartments in New York Product trainer and motivational speaker Software engineer + full stack developer (indeed, glassdoor) Conferences in ml + data science jobs
2.	Male-oriented	Finance job articles – CNBC businessinsider, emolument Data scientist – kdnuggets, glassdoor, datajobs, coursera Software development and software engineering Bollywood and small screen news Doctor jobs – naukri.com, Glassdoor, monsterindia, freshersworld Chef – catererglobal, indeed Gym Visiting career websites like IBM, Microsoft, Adobe,Google, Github for data science jobs Ml conferences
3.	Overlap between male and female gender	jobs of software engineer
4.	Gender Neutral	Event management – Naukri, indeed, LinkedIn When treatment comprises of both male and female interests Horoscope and interests pertaining to a culture Pinterest Alcohol support stress management Parenting Disability

5.3 Privacy

As the Google Ads Policy states , the privacy of user's information is taken into serious consideration whenever it is being used for improving the services[4]. Gender is one such information that is being pulled. When "Rather Not Say" is chosen, it is meant that the user does not want to reveal the gender. Yet predicting a gender by taking the online browsing behaviour into account contradicts the fact that the user wants to keep the gender private. Also, this predicted gender when shared with the third party websites may hamper the privacy of the user.

5.4 Preferences and bias

Preference can be described as having an inclination for an alternative over other while bias is prejudice for one alternative without any well established fact[9]. For website visits and web searches, all the data is preserved by Google as mentioned in Section 3.1. This data is modeled and analyzed by an algorithm which predicts demographics with reference to the web activity.

When a male user visits cooking, jewellery, makeup related websites, searches such keywords, he is predicted as "Female". This result, though inappropriate and prejudiced, won't effect one's day to day life in a significant manner. Whereas if one searches for career and finance related keywords, visits corresponding websites, and is predicted as "Male", it can be inferred that only males are thought to be eligible for big career opportunities. This also leads to denial of opportunities to many eligible candidates. Out of 21 jobs – for the web searches made, it was startling to discover that Google ads algorithm classifies more than fifty percent of the users searching for high paying jobs as "Male". Even web search for keywords like machine learning conference also predicts that the user is "Male" which can be termed as bias on basis of gender.

6 CONCLUSION

The results of this experiment indicate that there are privacy and bias issues associated with Google Ad Settings. Firstly, even when a Google user does not want to reveal the gender, the Google Ad algorithm predicts a conventional gender and it displays it in the Google Ad Personalization page. This is an indication of choice and privacy issues. Secondly, when it assigns a gender it is based on the websites searched, clicked or visited. This indicates that there is bias towards the websites a user visits – some websites might be male or female oriented while others are gender neutral. This indicates that there is a bias issue. We feel that this work will lead to a more comprehensive research on the algorithms used by Google Ads. An extended statistical analysis over a larger number of accounts may derive more generalized results and distinctions in the classification of websites which are leading to the prediction of gender.

This experiment can be conducted with a larger number of simulated users over a longer span of time to attain a better perception of what websites are more likely to identify the user as Male or Female. It can also be used to capture the difference between Male and Female online browsing behaviour as learned by the Google Ads algorithm.

REFERENCES

[1] 2019. Link to code and appendix,[Last visited on May 1, 2019]. https://github.com/nisha987/Google-gender
[2] Rachael Bennett. 2014. Google-Announcement for addtion of Rather Not Say and Custom gender options. https://plus.google.com/118279113645730324236/posts/FKK2trDERAC
[3] Amit Datta, Michael Carl Tschantz, and Anupam Datta. 2015. Automated experiments on ad privacy settings. *Proceedings on privacy enhancing technologies* 2015, 1 (2015), 92–112.
[4] Google. 2019. Making it easy to understand what data we collect and why, [Last visited on May 1, 2019]. https://safety.google/privacy/data/
[5] Google. 2019. Third-party sites apps with access to your account,[Last visited on May 1, 2019]. https://support.google.com/accounts/answer/3466521?hl=en
[6] Google. 2019. Verify your account,[Last visited on May 1, 2019]. https://support.google.com/accounts/answer/114129?hl=en&visit_id=636849961386601652-3618412408&rd=1
[7] Jonathan R Mayer and John C Mitchell. 2012. Third-party web tracking: Policy and technology. In *2012 IEEE Symposium on Security and Privacy*. IEEE, 413–427.
[8] Michael Carl Tschantz, Serge Egelman, Jaeyoung Choi, Nicholas Weaver, and Gerald Friedland. 2018. The accuracy of the demographic inferences shown on Google's Ad Settings. In *Proceedings of the 2018 Workshop on Privacy in the Electronic Society*. ACM, 33–41.
[9] WikiDiff. 2019. What is the difference between bias and preference?,[Last visited on May 1, 2019]. https://wikidiff.com/bias/preference
[10] Craig E Wills and Can Tatar. 2012. Understanding what they do with what they know. In *Proceedings of the 2012 ACM Workshop on Privacy in the Electronic Society*. ACM, 13–18.