Noticing the Other Gender on Google+

Diego Las Casas[1], Gabriel Magno[1], Evandro Cunha[1,2],
Marcos André Gonçalves[1], César Cambraia[2] and Virgilio Almeida[1]
[1]Dept. of Computer Science, [2]College of Letters – Federal University of Minas Gerais
Belo Horizonte, Minas Gerais, Brazil
{diegolascasas, magno, evandrocunha, mgoncalv, virgilio}@dcc.ufmg.br, nardelli@ufmg.br

ABSTRACT

Google+ provides a feature that has been overlooked in social media studies: the possibility of users setting their gender information not only as *female* or *male*, but as *other* instead. In this paper, we discuss this particularity and, more broadly, the issue of non-binary gender roles in the Web. By analyzing a large dataset, we characterize some aspects of self presentation, word use, network information and country of residence among users who choose different alternatives in the field *Gender*. On the whole, our main contributions are to present preliminary results and to shed light into the topic considered here – namely, the implications of having a third gender option to present oneself in an online social networking service.

Categories and Subject Descriptors

K.4 [**Computers and Society**]: Social Issues

Keywords

Online social networks; Gender issues; Google+

1. INTRODUCTION

Google+ is a major online social networking service launched by Google Inc. in June 2011. According to a recent report [14], Google+ grew in terms of active usage by 33% from June 2012 to May 2013 and was, in mid-2013, the second largest social platform in number of users, with about 360 million active profiles. As a comparison, at that time, Facebook, the most popular online social network (OSN), had almost 700 million active users; YouTube and Twitter, respectively the third and the fourth most popular ones, had slightly less than 300 million active users.

An interesting aspect of Google+ is that, during registration or any subsequent profile update, its users are given the option of setting their gender information not only as *female* or *male*, but also as *other*. This is an especially useful feature for profiles of members who do not associate with a particular gender – like, from one side, profiles of couples, groups and organizations; and from another side, profiles of users who identify themselves as, for instance, non-gender, transgender or intersex.

The main goals of this study are (a) to characterize the usage of an online social networking service by members who, for any reason, declare themselves as neither female nor male individuals, but as *other*; and (b) to raise awareness to the issue of non-binary gendered Web users. By investigating how users who declare themselves as other-gendered in Google+ behave in this network, we also provide contrastive information on the behavior of declaredly female and male members – and of Google+ users in general.

WebSci'14, June 23–26, 2014, Bloomington, IN, USA.

http://dx.doi.org/10.1145/2615569.2615692 .

2. BACKGROUND

In gender studies, the term *gender* is often used to refer to the social constructions of femininities and masculinities. It does not refer to biological differences – which are usually assigned to *sex* – but rather to cultural differences. In this section, we briefly review how gender is framed in OSNs, as well as how it has been studied by scholars.

2.1 Profile field "Gender" across networks

Until recently, Google+ was, among the most popular online social networking systems, the only one that provided to its users the option *other* as an alternative to *female* and *male* in the profile field *Gender*. Table 1 shows that, among the ten most popular OSNs and social curating websites in the beginning of 2014 [4], three did not provide a profile field *Gender* and only four provided an alternative to *female* and *male*. It is worth noticing that Pinterest and MySpace offered the *unspecified* alternative, which seems to give an idea of indefiniteness not given by *other*.

Table 1: Answer options for the field *Gender* during registration in the most popular OSNs

Online social network	Answer options for the field *Gender*	Is the answer compulsory?
Facebook	*Female / Male / Custom*	Yes
Twitter	Does not provide a field *Gender*	
LinkedIn	Does not provide a field *Gender*	
Google+	*Female / Male / Other*	Yes
Pinterest	*Female / Male / Unspecified*	Yes
Tumblr	Does not provide a field *Gender*	
Flickr	*Female / Male*	Yes
VK	*Female / Male*	Yes
Instagram	*Female / Male*	No
MySpace	*Female / Male / Unspecified*	Yes

There is a variety of reasons for a user choosing *other* as gender in an OSN. In the following lines, we will describe the three that we find most relevant for this study.

2.1.1 Not a person

As uses of social media are widely varied, many profiles can actually represent groups (e.g. bands), aggregations (e.g. couples), personas (e.g. fictional characters) or institutions (e.g. brands). All these kinds of profiles may eventually set *other* as gender.

In November 2011, Google+ added the possibility of creating special profile pages for institutions and organizations [5], which meant that these groups would not need to create new other-gendered profiles for this purpose. However, since a fraction of old business profiles did not enable this functionality, there are still organizations with profiles very similar to the ones of individual users.

2.1.2 Privacy worries

The option of setting gender information as private was enabled by Google+ in July 2011 [10] and those English-speaking users who set this information as private are identified by the pronoun *they* (as in *Alex updated **their** profile*), just like users who set their gender information as *other*[1].

Privacy-concerned users may declare their gender as *other* due to lack of knowledge of the feature of setting it as private, lack of interest in updating this field (if the profile was created before the pro-privacy feature was available) or because they do not want even Google and its commercial partners to have access to their gender identities.

2.1.3 Non-binarism

The *gender binary system* is a model of gender construction that classifies people into either *feminine* or *masculine*, *female* or *male*, *man* or *woman*, always as categorically distinct and mutually exclusive roles. It has been strongly criticized in the past decades as it is seen as a false dichotomy, since there are many more expressions of gender identities that cannot be collapsed into two discrete categories. Even the concept of *sex*, generally related to a biological basis, has its binarism called to question as it was shown to consist of numerous parts (e.g. chromossomal sex, anatomical sex, reproductive sex) and shades of gray (e.g. genital ambiguity and intersex conditions) [9]. Those who reject this binarism acknowledge *non-binary* or *transgender* identities. By allowing users to state that they are neither *female* nor *male*, Google+ opens up for non-binary gender identities to be expressed.

In February 2014, this subject made headlines due to the announcement that Facebook would allow users to pick a *custom gender* and it raised the discussion on the tailoring of online social media to transgender people [6].

2.2 On gender issues in social media

Two main approaches were identified by Van Doorn and Van Zoonen [13] as being taken by studies that address gender issues in digital environments: a) studies on *gender as identity* analyze the individual behavior of users of different genders either by enumerating gender differences in Internet usage or highlighting the importance of practices of experimentation of different gender roles in online situations; b) studies on *gender as social structure* describe gender inequalities, oppression or empowerment in technological settings. The authors noticed that both approaches frequently consider online and offline environments as separated entities, with one influencing the other but not the other way round. They suggest that gender roles should be approached as something that is both "shaping and shaped" by technology.

They also conclude that most studies show that the Internet reproduces inequalities and existing differences, but also enables new forms of transgressions of stereotypical gendered practices. This view is further explored by Van Doorn [12], who investigates micro-networks formed by close acquaintances in MySpace and concludes that these networks allow users to challenge standard norms of sexuality and gender roles through a cheerful interplay of stylistic forms.

Another work in this line is presented by De Ridder and Van Bauwel [3], who demonstrate how young users act their gender identities through commentaries in photos posted by other young users. They describe a rich symbolic environment in which gender roles are continually displaced and reaffirmed, showing tensions and contradictions that young users live in online interaction.

As boyd and Heer [1] frame it, social communication can be understood as performance, which is inherently embodied and contextually dependent. In this sense, *profiles* are digital bodies that interact to create social context within the network, which orients this performance. One of our goals is to explore empirically whether Google+ third gender option has had its effect in how users digital bodies interact in the online social environment.

3. DATA COLLECTION

To conduct this research, we collected Google+ public profiles of millions of users. Google+ members are able to choose the degree of visibility of the information available in their own profiles: social information, friends lists and posts published may be visible to friends, friends of friends, the general public or to customized lists of users, known as *circles*. For ethical and legal reasons, we only collected information set as public and did not attempt to access private information. Since our dataset was crawled from public Web pages, we can make it available upon request.

The data collection ran from March 23rd to June 1st, 2012. In order to retrieve the list of profiles to collect, we inspected the `robots.txt` file provided by Google+ and followed the corresponding sitemap to compile the lists of URLs of profiles. Because we collected the complete list of profiles provided by Google+, we believe that we collected information from all users with public profiles at the time of the data collection, gathering information from 160,304,954 profiles. By requesting the corresponding public friends list in the users' profiles, we also collected network information.

4. ANALYSES AND RESULTS

According to Magno et al. [7], *Gender* is the most shared profile field among those that can be set as private by Google+ users[2]: in our dataset, 126,531,842 (78.93%) of the more than 160 million users with profiles collected set their

[1]It is important to notice that the pronoun *they* is a common *preferred gender pronoun (PGP)* among transgenders and gender nonconforming people.

[2]*Name* is the most shared profile field, but users do not have the option of setting it as private.

gender information as public. Considering users who publicly shared their gender information, 80,683,714 (63.77%) declared themselves as *male*, 43,506,597 (34.38%) as *female* and 2,341,531 (1.85%) as *other*. In this section, we present the gender-related investigations performed in our dataset and the results obtained, as well as discuss findings.

4.1 Self presentation

Profile information is the first impression that users give to other network members. Here, we contrast profile information provided by Google+ users who report themselves as *female*, *male* and *other*, which should give insights into how they build their digital personas.

4.1.1 Looking for and Relationship status

Besides the field *Gender*, two Google+ profile fields have closed answer alternatives: *Looking for*, which suggests four alternatives, and *Relationship status*, which offers nine alternatives. In *Looking for*, network members can answer *friends*, *dating*, *relationship* and *networking*. Response to this field is not exclusive, meaning that users can choose from zero to all four alternatives. In our dataset, only 2,971,031 (1.9%) profiles filled this field in and set this information as public. In *Relationship status*, nine alternatives are available, going from traditional marital status (like *single* and *married*) to a joke option (*it's complicated*). This information is provided and set as public by 4,057,966 users (2.5% of the dataset).

Table 2 depicts information on the answers to these fields, showing the percentage of users who selected each option among those who made them available. Since members can select multiple options in the field *Looking for*, the sum of the percentages for each gender exceeds the value of 100.

Other-gendered users are less interested in *friends* (62.2% of other-gendered against 83.3% of females and 78.5% of males) and more interested in *networking* (62.5% of other-gendered against 40.2% of females and 55.7% of males) than the remaining users. This fact suggests that many of them are companies or organizations instead of individuals. For the remaining items, however, other-gendered users show interests in between of females and males.

Regarding relationship status, the large ratio of other-gendered users who reported their status as a non conventional or a non traditional one is meaningful: these users are the main selectors of the alternatives *it's complicated*, *in an open relationship*, *in a domestic partnership* and *in a civil union* – which together correspond to 33.0% of their responses, against only 8.0% of the responses of females and 7.6% of the responses of males.

4.1.2 Self-descriptive fields

Google+ profiles have three self-descriptive fields: *Tagline*, that invites users to shortly introduce themselves; *Introduction*, in which network members provide extended self-presentations; and *Bragging rights*, where they are encouraged to talk about their own reasons for pride.

Table 3 depicts the most frequent nouns in each of these fields by gender in order to exhibit common elements of self-presentation. The high prevalence of words related to business among other-gendered users (like *service*, *business* and *company*) reinforces our hypothesis that most users in this category are actually companies and commercial ventures. For this reason, we also provide the same information for users who not only report themselves as other-gendered, but publicly inform their relationship status – institutions and organizations will hardly fill this profile field in. We named this category as *other (V.R.)* (as in *visible relationship*) and Table 3 suggests that this filtering may be effective to help identifying individual users, as the most frequent words among them are similar to the ones of declared female and male members. This gives us a useful preliminary way of distinguishing between two kinds of other-gendered users, which we will use frequently along the paper.

Table 2: Information available in the fields *Looking for* and *Relationship status* by gender

Looking for	Total	Female	Male	Other
Friends	79.7%	83.2%	78.5%	62.2%
Networking	50.0%	40.2%	55.7%	62.5%
Relationship	25.3%	13.7%	32.3%	19.9%
Dating	22.3%	11.8%	28.7%	18.4%
Relationship status	**Total**	**Female**	**Male**	**Other**
Single	42.4%	35.4%	47.7%	24.0%
Married	27.2%	28.0%	26.7%	20.1%
In a relationship	17.3%	22.3%	13.8%	13.7%
Engaged	4.3%	5.4%	3.5%	4.5%
It's complicated	4.3%	4.7%	3.7%	16.3%
In an open relationship	2.0%	1.4%	2.3%	8.6%
In a domestic partnership	1.3%	1.4%	1.1%	3.7%
Widowed	0.7%	0.8%	0.6%	4.0%
In a civil union	0.5%	0.5%	0.5%	4.4%

Table 3: Most frequent nouns in each self-descriptive field by gender

Tagline				
Rank	Female	Male	Other	Other (V.R.)
1st	love	life	love	love
2nd	life	love	life	life
3rd	girl	man	music	fun
4th	fun	music	world	people
5th	world	guy	service	world
Introduction				
Rank	Female	Male	Other	Other (V.R.)
1st	love	love	service	love
2nd	life	life	business	life
3rd	years	years	people	people
4th	people	music	world	name
5th	name	time	company	time
Bragging rights				
Rank	Female	Male	Other	Other (V.R.)
1st	school	school	years	school
2nd	kids	kids	school	kids
3rd	children	years	world	years
4th	love	college	service	love
5th	years	life	business	life

4.2 Word use

Users communicate with others in Google+ through posts and comments. Since communication is gendered, users express their gender identities in their choice of words, as seen in a number of studies [8, 2]. We wished to understand how other-gendered users express themselves in relation to those with different gender identities: if they are conforming to a binary identity, they will tend to have an expression that is similar to one gender or another; otherwise, their expression will be in between the binaries.

For each gender, we calculated the frequency of usage of each word in the posts shared by its members and selected

the 1% most common words. Then, we found the percentage of male users who employed each word present in the $male + female$ frequency count and calculated the z-score of the percentages. This gave us a rough empirical measure representing how "masculine" a given word is in relation to an arbitrary "binary-only" basis: negative values mean words predominantly used by women (i.e., very feminine), while positive values indicate words employed mostly by men (i.e., very masculine). *User scores* were then obtained by averaging the score of all words employed by each user. This analysis was also focused in users filtered by the visibility of their relationship status, as described above.

Figure 1 shows the distribution of *user scores* by gender. There is a significant overlap between the distributions and, by construction, male users are shifted to the right, whereas female users are shifted to the left. This shows that our measure does differentiate between these two genders, which, however, are not completely separated – since they, as expected, do not have entirely different vocabularies. Other-gendered users are more spread than the other two genders and their distribution is centered between females and males. This hints to our hypothesis that their expression is less bounded by binary gender roles, as they seem to be able to transit between identities more associated with "masculine" and "feminine" performances.

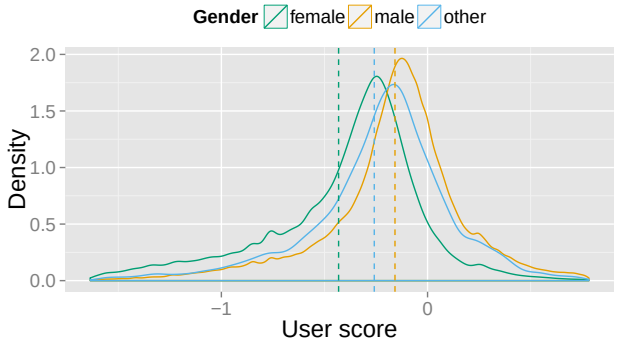

Figure 1: Density distribution of *user scores* by gender. Vertical Lines represent the averages.

4.3 Network structure

To undestand how other-gendered users are situated inside the network, we examined fundamental whole network and ego-network[3] metrics. Here, again, we also consider the visibility of the field *Relationship status*, since we observed that users who share more social information are likely to be more active(*e.g.* Users who share this field have higher average out-degree (46) than those who do not (19).

4.3.1 Clustering coefficient

The local clustering coefficient (CC) of an user is the probability of any two of its neighbors being neighbors themselves. Figure 2 depicts the mean value of CC for each gender. It shows that, independently of the visibility of the relationship status, other-gendered users have higher CC values than female and male users, which indicates that their ego-networks tend to be more densely connected.

[3]The ego-network of an user is formed by the user herself, her 1-hop neighbors and the respective connections among them.

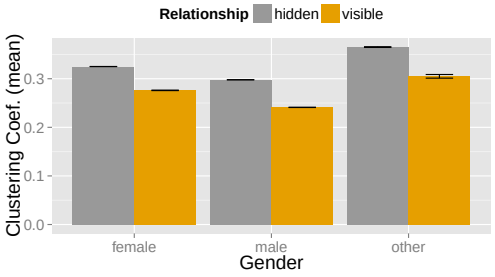

Figure 2: Mean clustering coefficient by gender ± standard errors

4.3.2 Homophily

In this analysis, we investigate the rate of homophily among gender groups in the network – that is, if users of a particular gender are more prone to connect to users of that same gender. For a given user, we calculated the fractions of neighbors of each gender, considering only neighbors with public gender information, so that the sum of the three fractions for a particular user is equal to 1.0. Although we present only the results for the outgoing edges, the incoming edges produce qualitatively equivalent results. Figure 3 shows the cumulative distribution functions of the three fractions for users with visible relationship information. We observe that users of a particular gender have higher fractions of neighbors of that same gender, indicating that there is, indeed, homophily in the network. Interestingly, values for all fractions of other-gendered members are in between those of female and male users, except for the fraction of neighbors of their own group.

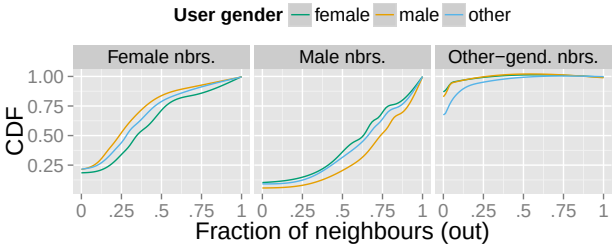

Figure 3: Cumulative distribution functions of fractions of neighbors by gender

Taken together, these results suggest that other-gendered users tend to be connected by a rich network with more other-gendered users, perhaps in a community fashion.

4.4 Country of residence as social context

Google+ has an international penetration, and thus its users are exposed to widely different social realities. Although describing the relationship that users from different societies have with gender identities is beyond our scope, we can investigate how expressive are other-gendered profiles across countries.

Users are able to supply a list of names of places where they have lived in the field *Places lived*. They can be as specific as they want and can describe a particular location

in distinct ways (e.g. "Rio de Janeiro", "Rio", "RJ") and levels of precision (e.g. "Paris", "Île-de-France", "France"). We extracted the geographic coordinates of the most recent information on the list – which is expected to indicate a user's current location – and translated them into valid location identifiers, which allowed us to identify the country of 22,578,898 (14.08%) users.

For each country in our dataset, we calculated the fraction of other-gendered users relative to the total amount of users. We call this measure *Other-Gendered Proportion* (OGP). To avoid biases and "mock locations"[4], we excluded from this analysis countries with less than 10,000 users in the dataset, which left us with a total of 21,563,667 users in 96 countries.

Figure 4 shows countries with highest and smallest OGPs. Some of them have extremely inexpressive OGPs, and it is worth noticing that some of the countries in the bottom 10 list are also countries with poor indicators of gender equality. In fact, a country's OGP seems to be related to gender equality indicators: the Spearman correlation between OGP and the Gender Inequality Index from 2012's Human Development Report [11] is -0.53, or -0.57 if we consider only users with visible relationship status (both values are statistically significant under $p < 0.0001$). Note that the correlations are negative because the indexes measure *inequality*.

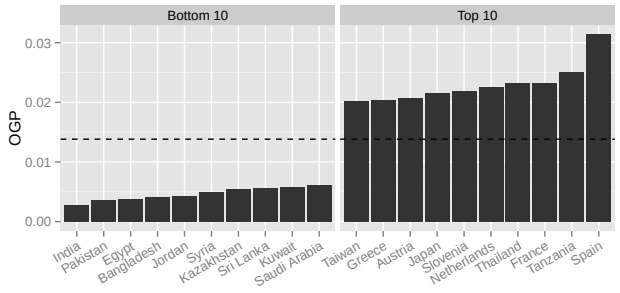

Figure 4: Country-wise proportion of users who declare themselves as *other-gendered*. The dashed horizontal bar indicates the average across countries

5. CONCLUDING REMARKS

In this paper, we analyzed users who chose *other* as their gender identity in Google+. We started by describing the profile information that they make public to show how they present themselves to the online world. We also related their word choices to those from binary-gendered users and showed that they are able to situate themselves in between the binary. By looking at their network features, we suggested that they link to additional other-gendered users and have, in general, dense ego-networks. We also observed that users from distinct countries have different dispositions towards expressing themselves as other-gendered, and we correlated this disposition to a social measure associated to gender equality. It should be noted that our results might be biased by the fact that the percentage of users from each gender who declare their own gender information can be asymmetric. Also, gender swapping is a known behavior in several online situations. These factors could not be controlled here and shall be addressed in future studies.

[4]Many users seem to state fake locations as a humorous tweak in their profiles.

Although *other* as a gender option impacts users' digital identities, its vagueness foreshadows the importance of OSNs allowing users to express their personal struggles through the network. The lack of a proper word for transgender identities hinders the clarity and profundity of this impact, as it unables users to distinguish among institutions, privacy-worried individuals and transgender users. Given the possibility, users will shape their profiles to better match their identities, and this will spread within the network.

Users tell stories by interacting with others in a (semi-) public environment. These stories are bounded by different media logics – such as an OSN user interface –, which in turn shape the representations that users build for themselves. The possibility of declaring oneself as *other* instead of collapsing to standard binarism allows network members to reshape their digital bodies, impacting gender identities and social contexts around them.

6. REFERENCES

[1] d. boyd and J. Heer. Profiles as conversation: Networked identity performance on Friendster. In *Proc. of Hawaii Int'l Conf. on System Sciences*, 2006.

[2] E. Cunha, G. Magno, M. A. Gonçalves, C. Cambraia, and V. Almeida. He votes or she votes? Female and male discursive strategies in Twitter political hashtags. *PLOS ONE*, 9(1):e87041, 2014.

[3] S. De Ridder and S. Van Bauwel. Commenting on pictures: teens negotiating gender and sexualities on social networking sites. *Sexualities*, 16(5-6), 2013.

[4] eBizMBA. Top 15 most popular social networking sites | January 2014. `http://bit.ly/1d2L2hr`, 2014.

[5] Google. Google+ pages: connect with all the things you care about. Google Blog, `http://bit.ly/1czGpxI`, November 2011.

[6] B. Griggs. Facebook goes beyond 'male' and 'female' with new gender options. CNN, `http://cnn.it/1l3zn9o`, February 2014.

[7] G. Magno, G. Comarela, D. Saez-Trumper, M. Cha, and V. Almeida. New kid on the block: exploring the Google+ social graph. In *Proceedings of the ACM Internet Measurement Conference (IMC'12)*, 2012.

[8] H. A. Schwartz, J. C. Eichstaedt, M. Kern, et al. Personality, gender, and age in the language of social media: The open-vocabulary approach. *PLOS ONE*, 8(9):e73791, 2013.

[9] S. Stryker and S. Whittle, editors. *The Transgender Studies Reader*. Routledge/Taylor & Francis, 2006.

[10] H. Tsukayama. Google Plus makes gender a private matter. The Washington Post, `http://wapo.st/1fMfmj9`, July 2011.

[11] United Nations Development Programme. Human development report. `http://hdr.undp.org`, 2012.

[12] N. Van Doorn. The ties that bind: the networked performance of gender, sexuality and friendship on MySpace. *New Media & Society*, 12(4), Nov. 2009.

[13] N. Van Doorn and L. Van Zoonen. Theorizing gender and the Internet: Past, present, and future. *The Routledge Handbook of Internet Politics*, 2008.

[14] T. Watkins. Suddenly, Google Plus is outpacing Twitter to become the world's second largest social network. `http://read.bi/1khuhED`, May 2013.