How Does Twitter User Behavior Vary Across Demographic Groups?

Zach Wood-Doughty*, **Michael Smith**†, **David A. Broniatowski**†, **Mark Dredze***
*Center for Language and Speech Processing
Johns Hopkins University, Baltimore, MD 21218
†Department of Engineering Management and Systems Engineering
George Washington University, Washington, DC 20052
zach@cs.jhu.edu, mikesmith@gwu.edu, broniatowski@gwu.edu, mdredze@cs.jhu.edu

Abstract

Demographically-tagged social media messages are a common source of data for computational social science. While these messages can indicate differences in beliefs and behaviors between demographic groups, we do not have a clear understanding of how different demographic groups use platforms such as Twitter. This paper presents a preliminary analysis of how groups' differing behaviors may confound analyses of the groups themselves. We analyzed one million Twitter users by first inferring demographic attributes, and then measuring several indicators of Twitter behavior. We find differences in these indicators across demographic groups, suggesting that there may be underlying differences in how different demographic groups use Twitter.

1 Introduction

Demographics have a central role in social science research, yet Twitter and other social media platforms often do not provide traditional demographic characteristics, such as age, gender and ethnicity. Inferring demographic attributes has thus been a frequent area of research (Burger et al., 2011; Pennacchiotti and Popescu, 2011; Volkova, 2015; Rao and Yarowsky, 2010; Mislove et al., 2011), enabling large-scale analysis of demographically identified social media posts. Demographic inference has been used in many Twitter analyses, including studies of mental health (Coppersmith et al., 2015), exercise (Dos Reis and Culotta, 2015), language (Eisenstein et al., 2011; Nguyen et al., 2013) and personality (Schwartz et al., 2013).

Several studies have examined the accuracy of demographic inference and the large-scale patterns it reveals. Chen et al. (2015) and Volkova et al. (2014) examined the effect of different types of information on the accuracy of demographic predictions. Mislove et al. (2011) examined how inferred demographics compare to known demographics outside of Twitter in the United States and measured in what ways the user-base of Twitter is biased compared to the population as a whole. Sloan et al. (2013) performed a similar analysis of gender and language among Twitter users in the United Kingdom.

However, even with accurate demographic inference tools, there may be other confounding factors that make it difficult to estimate variations of beliefs and behaviors across demographic groups. Since social media analysis relies on *how* people use platforms, variations in usage behaviors by different demographic groups could introduce biases in analyses and alter conclusions. For example, if one group tends to use Twitter nicknames more frequently, a name-based demographic classifier may make more errors on members of that group. Alternatively, if we use profile pictures to infer demographics and users of one demographic are less likely to share pictures of themselves, our results may under-represent that group. Pavalanathan and Eisenstein (2015) studied these issues for geolocation algorithms, finding that classifiers which infer users' locations identify a target population that differs from the general population of Twitter. A Pew Report survey indicated that social media users' privacy settings do vary across demographics, but did not look at specific behaviors (Madden, 2012).

This paper presents a first analysis of how differences in social media behaviors between demographic groups may confound demographic inference. Our aim is to identify potential sources of bias based on a large sample of Twitter users with demographic labels we infer using an ensemble of four classifiers for gender and ethnicity. We use systems that rely on several orthogonal sources of information to increase the robustness of our inference. We then measure various indicators of Twitter behaviors to identify potential differences across demographic groups. Our initial findings suggest that there may in fact be underlying differences in Twitter usage across these groups. This suggests that more work is needed to understand how these differences could impact the conclusions of Twitter analyses using inferred demographics.

2 Twitter User Data

We begin with a random sample of 5.4 million tweets taken from the 1% Twitter streaming API collected throughout the 12 months of 2016. From these tweets we sampled 1,000,000 users who had fewer than 500 followers and were not verified by Twitter, so as to exclude popular accounts, organizations, and "power users."

In May 2017, we attempted to download up to 200 of the most recent tweets of each user; this failed for the 18% of users who had made their accounts private or had deleted them altogether. For users who had tweeted fewer than 200 times, we retrieved their entire tweet history. This data reflects only those tweets that were publicly available at the time of our data collection. In total, we collected 158m tweets for 820k users, with a median of 200 tweets and a mean of 192 tweets per user that we could scrape.

3 User Behaviors

Our analyses focused on profile-based behaviors (invariant across all tweets) or those that could be estimated from (at most) 200 tweets. All behaviors appear in Table 1 in the order listed.

3.1 Profile Personalization

Many analyses of Twitter users are dependent on what information a user shares in his or her profile (Burger et al., 2011; Chen et al., 2015). We recorded whether each user included a custom profile image, URL, description, and location.

3.2 Temporal Information

To quantify each user's frequency of posting, we measured the average number of tweets per month from the time of account creation to the 2016 tweet.[1] We then computed the average of averages and the median average within each group. For the 38% of users who listed a timezone, we measured the normalized time-of-day of each tweet. Time-of-day data is useful for geolocation (Dredze et al., 2016) and understanding whether users are posting on Twitter from work or home.

3.3 Location Sharing

Several studies have examined location sharing behavior in Twitter (Mislove et al., 2011; Pavalanathan and Eisenstein, 2015; Dredze et al., 2013; Jurgens et al., 2015; Compton et al., 2014). However, these studies have not considered how this information may be correlated with demographic characteristics.

To determine the user's preference for sharing location information, we recorded whether a user had enabled geolocation sharing (a prerequisite for sharing GPS coordinates), and whether any of that user's tweets included GPS coordinates or a geotagged place. We also inferred locations for each tweet using Carmen (Dredze et al., 2013), a geolocation tool that estimates a user's location from the metadata from a single tweet. We recorded whether the Carmen tool could identify a country and/or a city from the user's profile.

3.4 User Interactions

Several previous studies have looked at how Twitter users interact with one another on the platform (Volkova and Bachrach, 2015; Bergsma et al., 2013; Volkova and Van Durme, 2015), including analyses of retweets (Luo et al., 2013; So et al., 2016; boyd et al., 2010) and replies or mentions (Honey and Herring, 2009; Hentschel et al., 2014).

For each user, we measured how many other users they mentioned across all tweets, how often they mentioned other users, how many of their tweets were retweets[2] or replies, and how often they shared images.

3.5 Devices

For each tweet, we record the contents of the "Source" field, which indicates from what type of device or platform the user posted. While there are many such platforms which represent hundreds of

[1] Tweet metadata includes date of account creation and total number of tweets from the account to date.

[2] We measure retweets via metadata, not the "RT" string.

different applications, we filter the results down to Android devices, iPhone devices, and desktop web clients. For each demographic group, we calculated the micro-averaged percent of tweets from each type of device and the macro-average of different types of devices used per user.

4 Demographic Classifiers

We used four separate approaches to infer the gender and ethnicity of the users in our dataset.

Demographer Demographer (Knowles et al., 2016) infers gender by first comparing a user's name against a namelist generated from the U.S. Social Security Administration, which includes the most likely gender. Second, for names not in the namelist, it uses an SVM to predict gender from character ngrams in the user's name.

Name RNN We extended Demographer by replacing the SVM with a recurrent neural network (RNN) which was trained on character sequences from Twitter names. We trained three models for predicting each of gender, 2-class ethnicity (Caucasian vs. African-American) and 3-class ethnicity (including Hispanic/Latino). As this classifiers was trained on the same data as the Demographer classifier, the two models had highly correlated predictions on users' genders.

Follower Lists Culotta et al. (2015) and Culotta et al. (2016) provide a model which uses a list of 1066 Twitter accounts which were highly correlated with demographic traits, according to Quantcast website data. The model predicts a user's gender and 4-class ethnicity (Caucasian, African-American, Hispanic/Latino, Asian) based on which, if any, of the Twitter accounts he or she follows. We gathered the entire list of followers for each of the 1066 Twitter accounts (totalling over 400 million users) to check which accounts were followed by which users. Because many users did not follow any of the accounts, this classifier did not always make a prediction.

Content Classifier Culotta et al. (2016) also provide a model that infers gender and 4-class ethnicity using the words in the user's tweet history. We ran this classifier on each of the users for which we could scrape a collection of tweets from 2017; because not all users mentioned terms within the model's vocabulary, it did not always make a prediction.

4.1 Comparing Demographic Classifiers

One issue in using this collection of classifiers is that they have different possible labels. The Follower Lists and Content Classifier methods include four categories for ethnicity, which does not match the number of categories from Demographer and Name RNN classifiers (two and three, respectively). For each classifier, White/Caucasian was the majority label in the training data and so the ambiguous instances may be classified as White. This is supported by the fact that 90% of our users were labeled as White by at least one classifier.

To account for the ethnicity label mismatch, we combine labels as follows: if the user was labeled as Asian by the Follower Lists or Content Classifier, we report the user as Asian; otherwise if the user was labeled as Hispanic/Latino by any classifier, we report that label; otherwise, if the user was labeled as Black/African-American by any classifier, we report that label; otherwise, if the user was labeled as White/Caucasian by two classifiers, we report that. This gives greater weight to ethnicity labels which could only be reported by a subset of the classifiers.[3] §5.1 discusses an alternative approach to handling this mismatch.

To reflect varying levels of agreement across the classifiers, we report separate numbers for how many classifiers agreed on gender. "M 2" means male according to two classifiers, which is a strict superset of "M 3", the users labeled as male by three classifiers. We ignored the 1.3% of users who were labeled as male by two classifiers and labeled as female by the other two classifiers.

5 Results

Table 1 shows results for gender and ethnicity, as well as the age of the user's account (discussed below). For many behaviors, there are marked differences across demographic groups. Across any two groups in the table (i.e. with at least 6.8% of the dataset per group), a macro-averaged difference of 1% between two proportions is statistically significant at the $p < 0.01$ level when using a two-tailed proportion test with

[3] There were 225k users twice-labeled as White/Caucasian which we reported as a different label on the basis of a single classifier. There were 107k users labeled as Black/African-American which we reported as Asian or Hispanic/Latino, and 9k users labeled as Hispanic/Latino which we reported as Asian.

a Bonferroni correction for 25 comparisons. Across the micro-averaged proportions for tweet percentages and time-of-day usage, a difference of 1% is significant using the same approach.

Gender There are several significant differences across inferred gender. Male-tagged users were significantly more likely to fill out the location and URL fields in their profiles, but were significantly less likely to enable geotagging.

There were only slight differences across time-of-day usage, though more male-tagged users had a timezone listed. Female-tagged users were more likely to use Android and iPhone devices, and less likely to use desktop web browsers or other sources.

Ethnicity Asian- and Hispanic/Latino-tagged users were far more likely to include a timezone in their profile, enable geotagging, share geotagged tweets, and include a location in their profile. Hispanic/Latino-tagged users had a higher proportion of tweets that were retweets, and were more likely to have a country identified by Carmen. White- and Black/African-American-tagged users had lower rates of almost all sharing-related behaviors, and were more likely to use iPhone devices and less likely to use Android devices or web clients.

Agreement as a Confounder Perhaps the most striking result is the difference between the gender groups with differing levels of classifier consensus ("M 2" vs. "M 3", and "F 2" vs. "F 3"). Users which had 3 classifiers in agreement for gender were significantly more likely to include a profile location or description.

This trend extends to the 2.0% of users for which all four gender classifiers agreed; the "F 4" and "M 4" users had significantly higher rates of almost every sharing behavior, including sharing one or more geotagged tweets (18.6% of users) and including a custom profile picture (99.0% of users). This indicates that agreement across classifiers is correlated with how much information a user is willing to share.

This is an important point, similar to that reported by Pavalanathan and Eisenstein (2015): propensity for sharing makes users easier to classify but presents a biased view of behavior. If correct, this may explain the differences between users labeled as either Asian or Hispanic/Latino compared to the overall usage rates. If our classifiers only report "Asian" when specific, rare indicators are present, it may be the case that users who create a profile with those indicators also share more information than the average user.

Account Age Another confound may come from how long a user has been been on Twitter, which could influence how much information they are willing to share. 50% of the users in our dataset created their account before October 9, 2014, which we used as the cutoff between "old" and "new" users. The final columns of Table 1 compares these two groups of users; there is a clear tendency for the old users to share more information in their profiles, but also to post far less frequently. Furthermore, we measured that among "F 2" and "M 2" users, 56.6% of users were old, whereas among "F 3" and "M 3" users, 72.4% were old. Among the 2.0% of users with unanimous gender classification, 85.6% were old. Thus, a user's account age is correlated with both how likely our classifiers are to agree upon a label, and how much information that user shares.

5.1 Limitations

An important limitation of our analysis is that not all ethnicity classifiers predict the same set of labels (§4.1). Only two classifiers label users as Asian, and only three classifiers label users as Hispanic/Latino. Because these classifiers were trained on different datasets with different ethnicity labels, we also don't know how correlated their predictions would be if they had all been trained on the same dataset. New training data could highlight correlations and differences between classifiers, and provide more evidence of convergent validity.

Furthermore, we only consider a small set of racial and ethnic groups. Our methods cannot label users as Native American or Pacific Islander, and there has been little to no work in identifying these groups in Twitter. Additionally, while Asian, Caucasian and Black are considered racial groups in traditional analysis, Hispanic/Latino descent is an ethnicity. Our classifiers conflate these distinctions; this issue and its implications for demographic surveys has been discussed in public health and social science research (Van den Berghe, 1978; Comstock et al., 2004; Gonzalez-Barrera and Lopez, 2015).

Finally, we do not have clear measurements of the precision and recall of each classifier, nor do

Behavior/Data	All	Gender F 2	Gender F 3	Gender M 2	Gender M 3	Ethnicity W	Ethnicity B	Ethnicity HL	Ethnicity A	Account Age O	Account Age N
% users in dataset	100	27.0	6.8	31.8	7.9	43.4	28.9	15.3	12.3	50.0	50.0
% users with tweets from 2017	82.0	81.9	81.8	81.9	82.0	82.0	82.0	81.9	82.0	81.9	82.0
% users with custom profile image	95.4	96.3	96.6	95.2	97.8	93.9	95.4	97.9	98.0	97.3	93.5
% users with profile URL	20.8	21.3	26.5	23.7	29.3	16.8	20.3	26.1	30.0	25.1	16.6
% users with profile description	78.0	76.1	81.0	77.0	80.7	74.1	79.1	80.7	85.3	81.0	75.0
% users with profile location	53.6	54.9	62.6	57.3	66.1	48.0	53.5	61.7	63.3	58.6	48.7
Average monthly tweets	739	673	432	696	413	806	775	481	735	391	1086
Median average monthly tweets	205	204	203	205	206	205	205	204	204	149	297
% users with timezone data	37.8	47.9	77.9	51.3	79.3	15.4	29.7	73.8	91.1	55.0	20.6
(m) % weekday tweets before 9am	20.7	19.7	18.2	20.0	17.4	20.0	20.8	17.7	24.0	18.7	26.2
(m) % weekday tweets 9am - 5pm	25.8	26.5	27.5	26.5	28.5	26.4	25.9	27.6	23.3	26.1	24.9
(m) % weekday tweets after 5pm	28.6	28.5	28.7	28.6	28.9	28.8	28.6	28.9	28.3	30.3	24.0
(m) % weekend tweets	24.9	25.3	25.7	25.0	25.2	24.9	24.7	25.8	24.4	25.0	24.9
% users with geotagging enabled	33.1	39.1	47.5	36.0	45.2	28.2	31.0	45.4	40.0	47.2	39.1
% users with 1+ geotagged tweet	7.9	10.8	15.5	10.0	14.9	6.1	6.8	13.0	10.8	11.4	4.5
% users with Carmen country	17.2	23.8	32.2	22.5	32.8	15.1	15.8	24.7	18.8	21.0	13.5
% users with Carmen city	8.6	11.7	16.2	11.9	18.4	7.6	8.2	11.5	9.6	11.1	6.2
Number of mentioned users per user	95	106	123	105	126	85	89	119	113	102	88
(m) % tweets that mention a user	22.3	23.0	24.5	24.7	28.7	22.6	21.8	22.5	22.4	23.7	20.8
(m) % tweets that are retweets	42.6	48.3	48.8	42.7	43.2	42.3	41.6	46.7	40.0	41.2	44.2
(m) % tweets that are replies	15.3	12.4	12.1	15.5	17.1	15.5	15.2	14.0	16.4	14.6	16.1
(m) % tweets that include an image	33.9	36.4	38.3	36.4	41.7	33.2	32.9	37.1	33.6	34.9	32.6
(m) % tweets from Android sources	30.5	32.0	30.0	30.3	27.8	28.8	28.7	36.6	30.9	27.2	34.6
(m) % tweets from iPhone sources	36.9	37.9	40.7	33.5	34.0	39.5	39.7	31.2	32.1	37.7	36.0
(m) % tweets from desktop web	9.0	9.4	10.4	11.5	15.5	7.4	7.5	12.4	12.2	9.7	8.2
Number of devices used per user	1.5	1.7	2.1	1.8	2.2	1.3	1.4	2.0	2.2	1.8	1.3

Table 1: **Behavior across groups.** For gender groups, 'M' stands for Male, 'F' for Female. '2' indicates that at least three gender classifiers agreed on the label; '3' indicates that all four did. For ethnicity groups, 'W' stands for White/Caucasian, 'B' for Black/African-American, 'HL' for Hispanic/Latino, and 'A' for Asian. For age (of account) groups, 'O' stands for old (user joined before Oct. 2014), 'N' for new. (m) indicates that a percent or average was computed via micro-averaging across users' tweets; all others are macro-averaged across users. Entries that require multiple tweets per user or timezone data are computed by ignoring the users for which that data is unavailable, which may introduce bias.

we know the distribution of the users for which our ensemble does not make a prediction. While we can identify some biases (e.g. the three-class ethnicity classifier biases against labeling users as Hispanic-Latino, due to limitations with the training data), there may be other systematic errors we cannot identify. Additional bias in our measurements could be introduced from the large proportions of our users for which we could not download tweets from 2017 and did not have a timezone. Better measurements of the performance of our classifiers would allow us to combine their predictions in a principled way to vary the agreement and accuracy of our ensemble, and validate the system's robustness.

6 Conclusion

We provide a preliminary look at possible confounds introduced by differences in how demographic groups use Twitter. We measure platform behaviors for a large set of Twitter users, and use recent tools to infer their demographic labels. Our analysis highlights several behavioral differences between groups that warrant further study. As demographic inference in social media becomes common practice, it is important to validate methodologies and test whether underlying biases exist. A "black-box" predictor that assumes all input fields are equally representative of the underlying population is likely to introduce biases against groups for which that assumption is false. We hope that future work can further examine such confounds to measure their effect on conclusions drawn in the social media analysis literature.

7 Acknowledgements

This work was supposed in part by the National Institute of General Medical Sciences under grant number 5R01GM114771.

References

Shane Bergsma, Mark Dredze, Benjamin Van Durme, Theresa Wilson, and David Yarowsky. 2013. Broadly improving user classification via communication-based name and location clustering on twitter. In *North American Chapter of the Association for Computational Linguistics (NAACL)*. pages 1010–1019.

danah boyd, Scott Golder, and Gilad Lotan. 2010. Tweet, tweet, retweet: Conversational aspects of retweeting on twitter. In *Hawaii International Conference on System Sciences (HICSS)*. IEEE, pages 1–10.

John D Burger, John Henderson, George Kim, and Guido Zarrella. 2011. Discriminating gender on twitter. In *Empirical Methods in Natural Language Processing (EMNLP)*. Association for Computational Linguistics, pages 1301–1309.

Xin Chen, Yu Wang, Eugene Agichtein, and Fusheng Wang. 2015. A comparative study of demographic attribute inference in twitter. In *International Conference on Weblogs and Social Media (ICWSM)*. pages 590–593.

Ryan Compton, David Jurgens, and David Allen. 2014. Geotagging one hundred million twitter accounts with total variation minimization. In *IEEE International Conference on Big Data*. pages 393–401.

R Dawn Comstock, Edward M Castillo, and Suzanne P Lindsay. 2004. Four-year review of the use of race and ethnicity in epidemiologic and public health research. *American journal of epidemiology* 159(6):611–619.

Glen Coppersmith, Mark Dredze, Craig Harman, and Kristy Hollingshead. 2015. From adhd to sad: analyzing the language of mental health on twitter through self-reported diagnoses. In *NAACL Workshop on Computational Linguistics and Clinical Psychology*.

Aron Culotta, Nirmal Ravi Kumar, and Jennifer Cutler. 2015. Predicting the demographics of twitter users from website traffic data. In *Conference on Artificial Intelligence (AAAI)*. pages 72–78.

Aron Culotta, Nirmal Ravi Kumar, and Jennifer Cutler. 2016. Predicting twitter user demographics using distant supervision from website traffic data. *J. Artif. Intell. Res.(JAIR)* 55:389–408.

Virgile Landeiro Dos Reis and Aron Culotta. 2015. Using matched samples to estimate the effects of exercise on mental health from twitter. In *Conference on Artificial Intelligence (AAAI)*. pages 182–188.

Mark Dredze, Miles Osborne, and Prabhanjan Kambadur. 2016. Geolocation for twitter: Timing matters. In *North American Chapter of the Association for Computational Linguistics (NAACL)*. pages 1064–1069.

Mark Dredze, Michael J Paul, Shane Bergsma, and Hieu Tran. 2013. Carmen: A twitter geolocation system with applications to public health. In *AAAI Workshop on Expanding the Boundaries of Health Informatics Using AI (HIAI)*.

Jacob Eisenstein, Noah A Smith, and Eric P Xing. 2011. Discovering sociolinguistic associations with structured sparsity. In *Association for Computational Linguistics (ACL)*. pages 1365–1374.

A Gonzalez-Barrera and MH Lopez. 2015. Is being hispanic a matter of race, ethnicity or both? *Pew Research Center* .

Martin Hentschel, Omar Alonso, Scott Counts, and Vasileios Kandylas. 2014. Finding users we trust: Scaling up verified twitter users using their communication patterns. In *International Conference on Weblogs and Social Media (ICWSM)*.

Courtenay Honey and Susan C Herring. 2009. Beyond microblogging: Conversation and collaboration via twitter. In *Hawaii International Conference on System Sciences (HICSS)*. IEEE, pages 1–10.

David Jurgens, Tyler Finethy, James McCorriston, Yi Tian Xu, and Derek Ruths. 2015. Geolocation prediction in twitter using social networks: A critical analysis and review of current practice. In *International Conference on Weblogs and Social Media (ICWSM)*. pages 188–197.

Rebecca Knowles, Josh Carroll, and Mark Dredze. 2016. Demographer: Extremely simple name demographics. In *EMNLP Workshop on Natural Language Processing and Computational Social Science*.

Zhunchen Luo, Miles Osborne, Jintao Tang, and Ting Wang. 2013. Who will retweet me?: finding retweeters in twitter. In *Conference on Research and development in information retrieval (SIGIR)*. pages 869–872.

Mary Madden. 2012. Privacy management on social media sites. *Pew Internet Report* pages 1–20.

Alan Mislove, Sune Lehmann, Yong-Yeol Ahn, Jukka-Pekka Onnela, and J Niels Rosenquist. 2011. Understanding the demographics of twitter users. In *International Conference on Weblogs and Social Media (ICWSM)*. volume 11.

Dong-Phuong Nguyen, Rilana Gravel, RB Trieschnigg, and Theo Meder. 2013. How old do you think i am? a study of language and age in twitter. In *International Conference on Weblogs and Social Media (ICWSM)*.

Umashanthi Pavalanathan and Jacob Eisenstein. 2015. Confounds and consequences in geotagged twitter data. In *Empirical Methods in Natural Language Processing (EMNLP)*.

Marco Pennacchiotti and Ana-Maria Popescu. 2011. A machine learning approach to twitter user classification. In *International Conference on Weblogs and Social Media (ICWSM)*. pages 281–288.

Delip Rao and David Yarowsky. 2010. Detecting latent user properties in social media. In *NIPS MLSN Workshop*.

H Andrew Schwartz, Johannes C Eichstaedt, Margaret L Kern, Lukasz Dziurzynski, Stephanie M Ramones, Megha Agrawal, Achal Shah, Michal Kosinski, David Stillwell, Martin EP Seligman, et al. 2013. Personality, gender, and age in the language of social media: The open-vocabulary approach. *PloS one* 8(9):e73791.

Luke Sloan, Jeffrey Morgan, William Housley, Matthew Williams, Adam Edwards, Pete Burnap, and Omer Rana. 2013. Knowing the tweeters: Deriving sociologically relevant demographics from twitter. *Sociological research online* 18(3):7.

Jiyeon So, Abby Prestin, Lyndon Lee, Yafei Wang, John Yen, and Wen-Ying Sylvia Chou. 2016. What do people like to share about obesity? a content analysis of frequent retweets about obesity on twitter. *Health communication* 31(2):193–206.

Pierre L Van den Berghe. 1978. Race and ethnicity: a sociobiological perspective. *Ethnic and racial studies* 1(4):401–411.

Svitlana Volkova. 2015. *Predicting Demographics and Affect in Social Networks*. Ph.D. thesis, Johns Hopkins University.

Svitlana Volkova and Yoram Bachrach. 2015. On predicting sociodemographic traits and emotions from communications in social networks and their implications to online self-disclosure. *Cyberpsychology, Behavior, and Social Networking* 18(12):726–736.

Svitlana Volkova, Glen Coppersmith, and Benjamin Van Durme. 2014. Inferring user political preferences from streaming communications. In *Association for Computational Linguistics (ACL)*. pages 186–196.

Svitlana Volkova and Benjamin Van Durme. 2015. Online bayesian models for personal analytics in social media. In *Conference on Artificial Intelligence (AAAI)*. pages 2325–2331.