Title: Self-Presentation and Information Disclosure on Twitter: Understanding Patterns and Mechanisms Along Demographic Lines

Authors: Nina Cesare, Emma Spiro, Tyler McCormick, and Hedwig Lee

ABSTRACT

Twitter data offers a wealth of opportunity for demographic research; however, one major challenge in this area is that individual-level characteristics of interest (e.g. age, race, gender) are not always made explicitly visible by users. Previous work has sought to find ways of extracting the demographic characteristics of Twitter users using metadata such as profile photos (McCormick, Lee, Cesare, and Shojaie, 2013), but many users choose not to display content that unambiguously discloses this information. In addition to this, while existing research has explored information disclosure and self-presentation within anonymous and "nonymous" social media sites (Turkle, 1995); little work has been done within semi-anonymous sites with highly variable, voluntary profile content, such as Twitter. Our goal for this project is to better understand patterns and mechanisms underlying missingness in information disclosure within Twitter user metadata by engaging in a mixed-methods analysis of users' profiles and behavior. In particular, we are interested in whether or not there are systematic differences in disclosure across demographic characteristics such as age, race and gender within this semi-anonymous site, and why these differences may exist. Our motivation for doing so is to provide a theoretical and methodological foundation for user-centric analyses of Twitter, as well as expand the scope of demographic research questions that can be addressed with this data source.

BACKGROUND

Understanding motivations behind implicit and explicit self-presentation within semi-anonymous social media spaces

When the internet first emerged as a space for interaction, researchers began to investigate how individuals used these spaces to express identity and create a "self" that has the potential to be either connected to or disconnected from their offline self depending on the affordances of the site. Early analyses of anonymous multi-user domains (MUDs) explored how the anonymity of online spaces impacted users' freedom of expression in constructing an online self (Turkle, 1995; Grazmuck, Martin and Zhao, 2009; Robinson, 2009). These studies found that under some conditions users choose to present themselves in truthful ways despite the anonymity of the site; under other conditions, however, the anonymity of MUDs allows users to explore hidden aspects of their identity or create and explore facets of themselves that do not exist offline (Turkle, 1995).

Research has also analyzed self-presentation strategies within "nonymous" social media spaces – sites where users are required to provide identifiable personal information and may establish connections with users who are familiar with their offline personas (Tufekci, 2008; Hogan, 2010; Lampe, Ellison and Steinfield, 2007). Examining displays of identity and information disclosure on Facebook within distinct ethno-racial groups, Grasmuck et al. (2009) found that users often bolster their ethno-racial identities by

manipulating their profiles to be similar to others with similar characteristics. In other words, users choose to display their offline characteristics both explicitly through the disclosure of truthful personal information and implicitly by crafting profiles similar to those of users with similar characteristics within "nonymous" sites.

Scholars have yet to deeply explore patterns of disclosure and self-presentation on semi-anonymous social media spaces. We define semi-anonymous social media spaces as sites in which users are allowed to present variable amounts of information about themselves within flexible fields. Users may establish connections with others who may or may not be familiar with their offline personas. Initial exploration of the data via our existing projects that seek to utilize demographic data embedded within Twitter user profiles (Lee, Cesare, McCormick, Morris and Shojaie, 2014; Cesare, Lee, McCormick, and Shojaie, 2014) has indicated that there is a remarkable amount of variation in the self-presentation strategies of Twitter users, and it is possible that this variation is systematically drawn along demographic lines.

Understanding Patterns of Disclosure

Social media data offers many exciting opportunities for demographic research, but it also presents many challenges. One of the major challenges of using this data is that individual-level characteristics of interest (e.g., age, race, gender) are not always made explicitly visible by users. Although past research has developed methods for extracting the demographic characteristics of Twitter users based on their profile photos (McCormick et al., 2013), many users choose not to display photos of themselves. Patterns of missingness within these data are not well understood. We do know, however, that the content of users' profiles is rich and may contain information beyond the profile photo that can cue the researcher (and followers of the profile) into the actual demographic characteristics of the user – including, but not limited to, externally hosted photos and content within tweets. Moreover, prior work suggests that behavioral patterns and information disclosure within Twitter could also be indicative of personal characteristics. (McKenna and Bargh, 1998; Florini, 2013; Sharma, 2013). Our objective in this work is to fill this gap in our understanding of disclosure online. Using a dataset based on behavioral traces and coded content of user profiles we will explore systematic differences in disclosure, then build a model to predict user demographic characteristics. . In this way, we will explore patterns of implicit and explicit disclosure of personal information and determine whether one can clue us in to the other. These developments will help fuel further research that focuses on characteristics of the users in addition to the content of their tweets, and will make Twitter more appealing as a source of data for social science research.

METHODS

Data Source

In this work we collect data from the population microblogging platform Twitter. Twitter allows users to record their thoughts in 140 characters or less. The text-based content of these messages may include personal updates, humor, or thoughts on media and politics. This concise format allows users to update their blogs multiple times per day, rather than every few days, as is the case with traditional blogging

platforms (Java et al 2007). Besides projecting their thoughts independently, users can communicate with one another either through private messages, by re-tweeting one another's tweets, or by using the *@reply* command. They may also contribute to broader conversations by including a *hashtag* identifier in their tweet. Tweets from those whom the user follows are displayed as a sequential feed that is updated in real time. User profiles are relatively sparse and contain, by default, the user's screen name. Users are also able to share a profile photo, their location, a name other than their screen name, a short description of themselves, a background or "cover photo", a personal website, and additional photos. These fields are not required.

Data Collection

In support of this research we gather profile and tweet information for a stratified sample of Twitter users. Utilizing three different methods for obtaining a close to random sample of Twitter users developed in prior work (Lee et al. 2014), we construct a stratified sample of Twitter users, based on whether or not they disclose their race in their Twitter profile picture. We use the Twitter Application Programming Interface (API) to obtain individual-level covariates about each user, such as their profile picture, account creation date and number of the social ties, as well as a set of messages posted by each user. We use this dataset as the starting point for our mixed-methods approach to exploring disclosure and missingness within Twitter user metadata.

Content Analysis of User Profiles

We perform a content analysis of all users in the sample, coding elements of the profile to build an augmented dataset with comprehensive information on the self-presentation strategies of the users sampled. Basic coding will include: noting the presence and absence of personal identifiers within users' Twitter profiles (i.e. whether users choose to display a profile photo, whether they choose to share their location, whether they include their real name within their profiles, whether they link to other social media outlets under their name). Beyond this, we will develop an open coding scheme to provide a more detailed depiction of trends in what people choose to share within available fields. For example, among profile photos we may choose to distinguish between users who display an identifiable photo of their own face, those who choose to present themselves within a large group, or those who choose to display a photo of an animal or company logo. We may also note whether the profile contains references to race in its description or username. As part of this component of our analysis, we will determine whether there are systematic differences in regards to self-presentation styles on Twitter and whether these patterns in self-presentation are divided along demographic lines.

The content analysis component of our work provides important descriptive information about general trends and systematic differences in disclosure on Twitter. However, there are likely individuals who provide little to no demographic information about themselves in any of their profile data or may deliberately provide misinformation. This missingness of Twitter user metadata is a significant barrier to using this data source for demographic research. We address this challenge in two ways. First, we explore automated techniques for predicting demographic characteristics based on other information about users as a means of better understanding patterns of disclosure and missingness, and second, we

survey Twitter users about their demographic characteristics, disclosure choices and the motivation for these choices.

Addressing Mechanisms that Drive Information Disclosure Within Twitter

In addition to examining patterns of missingness within these data in a way that increases the usability of these data for demographic research, we are also interested in exploring the mechanisms that drive self-presentation within the context of Twitter. Specifically, we are interested in learning about whether individuals' choices regarding the content and style of their personal information disclosure is linked to their demographic characteristics. We may wish to know, for example, if the decision to display or not display a profile photo is believed by the user to be a the same strategy used by others within the same age, race or gender category. If this is the case, then this decision can be considered a means of expressing an identity linked to these characteristics. We plan to design a short survey that seeks to capture these trends in user preferences. Using a targeted sampling strategy, we will recruit individual Twitter users from each of our stratified sample groups to participate in the survey.

Addressing Patterns in Information Disclosure on Twitter

In this component of our project, we explore strategies for dealing with missingness in demographic disclosure on Twitter. We use a rich dataset of coded Twitter profiles, as described previously, to build a predictive model of demographic characteristics such as age, race and gender, that utilized other information known about individuals, such as the posting behavior, commonly used words, etc. Doing so will provide information on how disclosure and missingness are patterned within the data along demographic lines. Our goal in this work is to build a model that will allow researchers to impute demographic characteristics to reduce the effect of missingness in social media data, potentially making social media dataset more appropriate to addressing core demographic research questions.

REFERENCES

Cesare, Nina, Hedwig Lee, Tyler McCormick, Ali Shojaie. 2014. “Examining Demographic Trends in Public Opinion on Twitter: The Case of Gun Control.” Population Association of American 2014 Annual Meeting. Boston, MA.

Florini, Sarah. 2013. “Tweets, Tweeps, and Signifyin’: Communication and Cultural Performance on ‘Black Twitter.’” *Television & New Media* 15(3):223–37.

Grasmuck, Sherri, Jason Martin, and Shanyang Zhao. 2009. “Ethno-Racial Identity Displays on Facebook.” *Journal of Computer-Mediated Communication* 15(1):158–88.

Hogan, B. 2010. “The Presentation of Self in the Age of Social Media: Distinguishing Performances and Exhibitions Online.” *Bulletin of Science, Technology & Society* 30(6):377–86.

Java, Akshay, Xiaodan Song, Tim Finin, and Belle Tseng. 2007. "Why we twitter: understanding microblogging usage and communities." Pp. 56-65 in *Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 workshop on Web mining and social network analysis*. San Jose, California: ACM

Lampe, CAC, Nicole Ellison, and Charles Steinfield. 2007. “A Familiar Face (book): Profile Elements as Signals in an Online Social Network.” *Proceedings of the SIGCHI …* 435–44.

Lee, Hedwig, Nina Cesare, Tyler McCormick, Julia Morris, Ali Shojaie. 2014. “Redrawing the “Color Line”: Examining Racial Homophily of Associative Networks in Social Media.” Population Association of American 2014 Annual Meeting. Boston, MA.

McCormick, Tyler, Hedwig Lee, Nina Cesare and Ali Shojaie. 2013. “Using Twitter for Demographic and Social Science Research: Tools for Data Collection.” CSSS Working Paper 127. University of Washington Center for Statistics and the Social Sciences, Seattle, WA.

Mckenna, Katelyn Y. A., and John A. Bargh. 1998. “Coming Out in the Age of the Intemet : Identity ‘ Demarginalization ’ Through Virtual Group Participation.” *Journal of Personality and Social Psychology* 75(3):681–94.

Robinson, L. 2007. “The Cyberself: The Self-Ing Project Goes Online, Symbolic Interaction in the Digital Age.” *New Media & Society* 9(1):93–110. Retrieved June 16, 2014 (http://nms.sagepub.com/cgi/doi/10.1177/1461444807072216).

Sharma, Sanjay. 2013. “Black Twitter? Racial Hashtags, Networks and Contagion.” *New Formations: A Journal of Culture/Theory/Politics* 78(1):46–64.

Tufekci, Z. 2007. “Can You See Me Now? Audience and Disclosure Regulation in Online Social Network Sites.” *Bulletin of Science, Technology & Society* 28(1):20–36.

Turkle, Shelly. 1995. “Life on the Screen: Identity in the Age of the Internet.” New York: Simon and Schuster.