CS 6474/CS 4803 Social Computing:
Social Computing Theories: Identity

Munmun De Choudhury

munmund@gatech.edu

Week 4 | February 4, 2021

"What's in a name? That which we call a rose By any other name would smell as sweet."

In Romeo and Juliet (II, ii, 1-2), Juliet tells Romeo that a name is an artificial and meaningless convention, and that she loves the person who is called "Montague", not the Montague name and not the Montague family.

Identity and Deception in the Virtual Community

Summary

- Identity in the Usenet environment (newsgroups).

- Models of honesty and deception

  - Assessment signals: They are costly and the cost is directly related to the trait being advertised.

    - Stag's horns

  - Conventional signals: Here, the signal is correlated with a trait is by custom or convention: the sender need not possess the trait in order to make the signal.

    - Powerlifter tshirt

- Assessment signals are reliable, since sending an assessment signal requires that the sender possess the relevant trait.

- Conventional signals are not reliable.

- Identity in voice and signal.

- "Signatures" to indicate identity.

Did you come up with interesting examples of conventional and assessment signals? [Ex. Online dating sites]

Class Exercise 1a

Class Exercise 1b

There are also consequences and implications of different identity choices...

Gender identity and lexical variation in social media[1]

David Bamman,[a] Jacob Eisenstein[b] and Tyler Schnoebelen[c]

a. Carnegie Mellon University, Pennsylvania
b. Georgia Institute of Technology, Georgia
c. Idibon, Inc., California

We present a study of the relationship between gender, linguistic style, and social networks, using a novel corpus of 14,000 Twitter users. Prior quantitative work on gender often treats this social variable as a female/male binary; we argue for a more nuanced approach. By clustering Twitter users, we find a natural decomposition of the dataset into various styles and topical interests. Many clusters have strong gender orientations, but their use of linguistic resources sometimes directly conflicts with the population-level language statistics. We view these clusters as a more accurate reflection of the multifaceted nature of gendered language styles. Previous corpus-based work has also had little to say about individuals whose linguistic styles defy population-level gender patterns. To identify such individuals, we train a statistical classifier, and measure the classifier confidence for each individual in the dataset. Examining individuals whose language does not match the classifier's model for their gender, we find that they have social networks that include significantly fewer same-gender social connections and that, in general, social network homophily is correlated with the use of same-gender language markers. Pairing computational methods and social theory thus offers a new perspective on how gender emerges as individuals position themselves relative to audiences, topics, and mainstream gender norms.

LGBT Parents and Social Media: Advocacy, Privacy, and Disclosure during Shifting Social Movements

Lindsay Blackwell, Jean Hardy, Tawfiq Ammari, Tiffany Veinot, Cliff Lampe, Sarita Schoenebeck
School of Information
University of Michigan
{lblackw,jkhardy,tawfiqam,tveinot,cacl,yardi}@umich.edu

ABSTRACT

Increasing numbers of American parents identify as lesbian, gay, bisexual, or transgender (LGBT). Shifting social movements are beginning to achieve greater recognition for LGBT parents and more rights for their families; however, LGBT parents still experience stigma and judgment in a variety of social contexts. We interviewed 28 LGBT parents to investigate how they navigate their online environments in light of these societal shifts. We find that 1) LGBT parents use social media sites to detect disapproval and identify allies within their social networks; 2) LGBT parents become what we call *incidental advocates*, when everyday social media posts are perceived as advocacy work even when not intended as such; and 3) for LGBT parents, privacy is a complex and collective responsibility, shared with children, partners, and families. We consider the complexities of LGBT parents' online disclosures in the context of shifting social movements and discuss the importance of supporting individual and collective privacy boundaries in these contexts.

AUTHOR KEYWORDS

LGBT; parents; families; advocacy; privacy; social media.

recognition of LGBT families, with or without children, is growing: in 2014, 55% of Americans polled were supportive of same-sex marriage, compared with only 35% in 2001 [61].

This change coincides with broader social movements characterized by significant and cumulative successes in LGBT peoples' fights for social and political inclusion in the U.S. and in many other countries. In 2015 alone, a number of key events highlighted these movements: on February 18, Kate Brown became the first openly bisexual U.S. governor [17]. On March 31, Tokyo's Shibuya ward became the first region in eastern Asia to recognize same-sex marriage [30]. On May 23, Ireland became the first nation to approve same-sex marriage by popular vote [73]. On June 1, Caitlyn Jenner became the first openly transgender woman to be featured on the cover of *Vanity Fair* [7]. Most significantly in the U.S., on June 26, 2015, the Supreme Court ruled state-level bans on same-sex marriage to be unconstitutional [72].

Despite these high-profile events, discrimination toward LGBT individuals remains a serious problem in the U.S. [21,74–77]. ENDA, the Employee Non-Discrimination Act,

Digital Footprints and Changing Networks During Online Identity Transitions

Oliver L. Haimson	**Jed R. Brubaker**	**Lynn Dombrowski**	**Gillian R. Hayes**
Dept. of Informatics	Dept. of Info. Science	School of Informatics	Dept. of Informatics
Univ. of California, Irvine	Univ. of Colo., Boulder	and Computing, IUPUI	Univ. of California, Irvine
Irvine, CA, USA	Boulder, CO, USA	Indianapolis, IN, USA	Irvine, CA, USA
ohaimson@uci.edu	jed.brubaker@colorado.edu	lsdombro@iupui.edu	gillianrh@ics.uci.edu

ABSTRACT

Digital artifacts on social media can challenge individuals during identity transitions, particularly those who prefer to delete, separate from, or hide data that are representative of a past identity. This work investigates concerns and practices reported by transgender people who transitioned while active on Facebook. We analyze open-ended survey responses from 283 participants, highlighting types of data considered problematic when separating oneself from a past identity, and challenges and strategies people engage in when managing personal data in a networked environment. We find that people shape their digital footprints in two ways: by editing the self-presentational data that is representative of a prior identity, and by managing the configuration of people who have access to that self-presentation. We outline the challenging interplay between shifting identities, social networks, and the data that suture them together. We apply these results to a discussion of the complexities of managing and forgetting the digital past.

times, this connection to one's history through data traces can be appealing. Many people enjoy looking at old digital artifacts and reminiscing about, for instance, the time before they had children, or the period when they dyed their hair red. However, for others, data from the past can be reminders of difficult times and identities they would rather forget or have others forget. Identity transitions are, for some, a substantial move away from a difficult past identity, a move that can be made more difficult by the persistence of *digital footprints*: the trails and artifacts that people leave behind when interacting in a digital setting [37]. Digital footprints link the past with the present in ways that may be problematic during identity transitions.

In this work, we describe how people manage digital artifacts during gender transition on Facebook. Our analysis contributes a deep understanding of how people manage digital artifacts and online social relationships during identity changes. How do social network sites (SNSs) enable and inhibit networked presentations of self? Studying gen-

"Too Gay for Facebook": Presenting LGBTQ+ Identity Throughout the Personal Social Media Ecosystem

MICHAEL A. DEVITO, Media, Technology, and Society, Northwestern University, USA
ASHLEY MARIE WALKER, Media, Technology, and Society, Northwestern University, USA
JEREMY BIRNHOLTZ, Communication Studies, Northwestern University, USA

Most US social media users engage regularly with multiple platforms. For LGBTQ+ people, this means making self-presentation decisions not just on one platform, but many. These choices are made in the face of sometimes-overlapping platform environments, which can have consequentially different norms, audiences, and affordances. Moreover, many LGBTQ+ users face high stakes in online self-presentation, due to the risk of stigmatization of their LGBTQ+ identity, increasing the importance of self-presentation decisions that enable them to achieve their goals and avoid stigmatization. This combination of environmental complexity and high stakes is not adequately accounted for in existing work on self-presentation, but doing so is important to support and understand the experiences of LGBTQ+ and other potentially stigmatized users. We adopt an ecological approach to an interview and cognitive mapping study of 20 LGBTQ+ social media users. We find that participants employ the platforms, audiences, affordances, and norms within what we call their "personal social media ecosystems" to avoid stigmatization while still allowing for expression of their LGBTQ+ identity and the flexibility to adjust their presentation over time.

CCS Concepts: • **Human-centered computing: Empirical studies in collaborative and social computing** • Human-centered computing: Empirical studies in HCI • **Social and professional topics: Sexual orientation**

KEYWORDS

Sexual and gender minorities; LGBTQ; LGB; queer; self-presentation; social media; technology ecosystems; identity management; disclosure; privacy

4chan and /b/: An Analysis of Anonymity and Ephemerality in a Large Online Community

What is anonymity?

- Security researchers define anonymity as unidentifiability "within a set of subjects".

- In sociology, Gary Marx's analysis: *being anonymous means a person cannot be identified according to any of seven dimensions of identity knowledge, that is, the person's* legal name, location, pseudonyms *that can be* linked to the person's legal name or location, pseudonyms *that cannot be linked to specific identity information but that* provide other clues to identity, *revealing patterns of* behavior, *membership in a* social group, *or information, items, or skills that indicate* personal characteristics.

4chan and /b/

- An English-language imageboard website.

- Users generally post anonymously, with the most recent posts appearing above the rest.

- 4chan is split into various boards with their own specific content and guidelines.

- Registration is not possible

- /b/ permits discussion and posting of any sort of content

- /b/ is the only board in 4chan where anyone can post grotesque material.

- /b/ is among the boards on 4chan which has a NSFW designation

Summary

- Study of 4chan's /b/

- Contributions:

  - Dominated by playful exchanges of images and links

  - Most threads spend just five seconds on the first page and less than five minutes on the site before expiring

  - 90% of posts are made by fully anonymous users, with other identity signals adopted and discarded at will

/b/ users also adopt various textual, linguistic, and visual cues to signal social status, e.g., in-dialect writing (/b/ slangs) serves as an entry-level signal of membership and status.

What other methods are adopted in other communities for the same?

Do you see such signaling only in anonymous sites?

Class Activity 2

Why do people seek anonymity online?
[People are rarely anonymous in the offline world, unless they are a criminal or a CIA agent!]

Why Do People Seek Anonymity on the Internet? Informing Policy and Design

Ruogu Kang[1], Stephanie Brown[2], Sara Kiesler[1]
Human Computer Interaction Institute[1]
Department of Psychology[2]
Carnegie Mellon University
5000 Forbes Ave., Pittsburgh, PA 15213
ruoguk@cs.cmu.edu, smb1@andrew.cmu.edu, kiesler@cs.cmu.edu

ABSTRACT

In this research we set out to discover why and how people seek anonymity in their online interactions. Our goal is to inform policy and the design of future Internet architecture and applications. We interviewed 44 people from America, Asia, Europe, and Africa who had sought anonymity and asked them about their experiences. A key finding of our research is the very large variation in interviewees' past experiences and life situations leading them to seek anonymity, and how they tried to achieve it. Our results suggest implications for the design of online communities, challenges for policy, and ways to improve anonymity tools and educate users about the different routes and threats to anonymity on the Internet.

literature that exists mainly derives from studies of one or a few online communities or activities (e.g., the study of 4chan in [5]). We lack a full understanding of the real life circumstances surrounding people's experiences of seeking anonymity and their feelings about the tradeoffs between anonymity and identifiability. A main purpose of the research reported here was to learn more about how people think about online anonymity and why they seek it. More specifically, we wanted to capture a broad slice of user activities and experiences from people who have actually sought anonymity, to investigate their experiences, and to understand their attitudes about anonymous and identified communication.

Another purpose of this research was to understand the

Why do people seek anonymity online?

Category	Advantages of being anonymous	Advantages of being identified
Social connections	Avoid disliked others Avoid commitment to the community Lower barrier to new relationships Protect others one cares about	Connect to real life friends Have stronger social connections Encourages more participation
Reputation and trust	Give honest rating/ recommendation	Good for reputation building Gain trust from other users
Image building	Have control over personal image Avoid embarrassment /judgment /criticism	Avoid harsh criticism Consistent with self-image
Emotional benefit	Feel relax and comfortable Feel cool and sophisticated	Feel real, integrated Feel closer to people
Express opinion	Feel free to express views	Avoid irresponsible behavior
Privacy	Have more control over personal information disclosure	Look innocent
Security	Protect personal safety Avoid legal repercussion/spam/stalk/lost of property	Hide in the crowd
Ease of use	Saves effort to log in	Easy to remember account

Table 2. Perceived tradeoffs of being anonymous vs. being identified

Kang et al. 2013

Identity Management and Mental Health Discourse in Social Media

Umashanthi Pavalanathan
School of Interactive Computing
Georgia Institute of Technology
Atlanta, GA 30308
umashanthi@gatech.edu

Munmun De Choudhury
School of Interactive Computing
Georgia Institute of Technology
Atlanta, GA 30308
munmund@gatech.edu

ABSTRACT

Social media is increasingly being adopted in health discourse. We examine the role played by identity in supporting discourse on socially stigmatized conditions. Specifically, we focus on mental health communities on reddit. We investigate the characteristics of mental health discourse manifested through reddit's characteristic 'throwaway' accounts, which are used as proxies of anonymity. For the purpose, we propose affective, cognitive, social, and linguistic style measures, drawing from literature in psychology. We observe that mental health discourse from throwaways is considerably disinhibiting and exhibits increased negativity, cognitive bias and self-attentional focus, and lowered self-esteem. Throwaways also seem to be six times more prevalent as an identity choice on mental health forums, compared to other reddit communities. We discuss the implications of our work in guiding mental health interventions, and in the design of online communities that can better cater to the needs of vulnerable populations. We conclude with thoughts on the role of identity manifestation on social media in behavioral therapy.

Categories and Subject Descriptors

J.4. [**Social and Behavioral Sciences**]: Psychology

The nature of online mental health discourse, however, may vary depending on the *nature of identity* adopted by an individual. This is likely to be particularly valid in the case of mental illness, since it is considered socially stigmatic [8]. Literature in sociology also supports this observation. In his celebrated book "Stigma" [14], Goffman examined how, individuals with a socially discredited attribute such as mental illness, tend to manage impressions of themselves in social settings—in order to protect their identities. However we note that in online settings, such as on social media, this constraint may be circumvented. This is because individuals may choose to withhold their actual identities allowing themselves to engage in more candid self-disclosure than is possible in offline settings, or through their identified online personas.

Our motivation for this research is also rooted in the rich literature on online identity construction, which has been recognized as a key aspect of online communities [11, 28]. Prior work demonstrates dissociative anonymity (a resistance to attach to offline identity or to their actual account/online persona), for instance, can be the foundation of online disinhibition [24]. Online disinhibition, the ability to avoid being "visible, verifiable, and accountable", leads people to act differently than they would in offline settings [6]. Social media naturally provides us with a rich ecosystem where we can study ways in which individuals manage their identities to engage in discourse on a stigmatized condition like mental

The throwaway cohort hints at a sense of urgency or desire to act: ("just need to", "to do something", "am going to"): *now i'm not crazy, i'm not a danger to any one, i just need to stay busy until i can see a new therapist in the next couple of days.*

Throwaway posts extensively share posters' personal beliefs and fear. This might reveal their vital constructs and private, sensitive informational attributes ("if I could", "part of me", "because I know").

The throwaway cohort also expresses a desire to avail help/need from the community ("want to talk", "what do i"): *i think about suicide at least once or twice a day but im not sure if i could go through with it.*

Category	Throw.	Identified	t-stat	p
Cognitive Attributes				
negation	0.0309	0.0282	5.537	***
certainty	0.0171	0.0160	3.561	**
Linguistic Style Attributes				
Lexical Density				
verbs	0.1899	0.1847	4.872	***
nouns	0.1777	0.1893	-5.396	***
Temporal References				
Social/Personal Concerns				
family	0.0055	0.0046	4.948	***
health	0.0138	0.0155	-3.448	**
Interpersonal Awareness				
1st person singular	0.1180	0.1123	6.141	***
2nd person	0.0045	0.0069	-6.277	***
3rd person	0.0195	0.0175	3.733	**
Social Attributes				
post length	332.0886	253.5124	13.361	***
median comment length	78.5943	64.9318	8.0622	***

Table 6: **Results of independent sample t-test between throwaway mental health posts and identified mental health posts.**

The Secret Life of Online Moms:
Anonymity and Disinhibition on YouBeMom.com

Sarita Yardi Schoenebeck

School of Information, University of Michigan

yardi@umich.edu

Abstract

Moms are one of the fastest growing demographics online. While much is known about where they spend their time, little is known about how they spend it. Using a dataset of over 51 million posts and comments from the website YouBeMom.com, this paper explores what kinds of topics moms talk about when they are not constrained by norms and expectations of face-to-face culture. Results show that almost 5% of posts are about *dh*, or "dear husband," but these posts tend to express more negative emotion than other posts. The average post is only 124 characters long and family and daily life are common categories of posting. This suggests that YouBeMom is used as a fast-paced social outlet that may not be available to moms in other parts of their lives. This work concludes with a discussion of anonymity and disinhibition and puts forth a new provocation that moms, too, spend time online "for the lulz."

Prior work has examined how moms seek health information and social support online (Plantin and Daneback 2009; Sarkadi and Bremberg 2005). Related work has also studied the role of the Internet in family life (Boneva et al. 2004; Kraut et al. 2002; Mesch 2006). Though motherhood and the culture of information sharing has been studied extensively in offline settings (e.g. Scott, Brady, and Glynn 2001), less work has focused on how moms spend their time online. This is an important demographic to study. One-third of all bloggers are moms, older moms are one of the fastest growing demographics on Facebook, and younger moms are 85% more likely to visit Facebook than the average user (Nielsen 2009).

What moms talk about when they are not constrained by norms and expectations of face-to-face culture

In some posts, **dh** is used in an affectionate way that implies that the husband is in fact dear to the poster. For others, **dh** is used cynically, often in the context of a deeply sarcastic or angry post:

*"If your **dh** treated you badly during your high risk pg [pregnancy], and got into a fight at the hospital after **db** was born, didn't show up the next day, would that be enough for you to divorce?"*

An interesting case of online anonymity: "*for the lulz*"

Schoenebeck 2014

Is anonymity a binary attribute?

Whisper

- Whisper was a mobile app that allowed users to post and reply to public messages on top of an image (e.g. Internet memes), all using anonymous user identifiers.

- Whisper did not associate any personal identifiable information with user IDs, does not archive any user history, and does not support persistent social links between users

The Many Shades of Anonymity:
Characterizing Anonymous Social Media Content

Denzil Correa[†], Leandro Araújo Silva[‡], Mainack Mondal[†],
Fabrício Benevenuto[‡], Krishna P. Gummadi[†]
[†] Max Planck Institute for Software Systems (MPI-SWS), Germany
[‡] Federal University of Minas Gerais (UFMG), Brazil

Abstract

Recently, there has been a significant increase in the popularity of anonymous social media sites like Whisper and Secret. Unlike traditional social media sites like Facebook and Twitter, posts on anonymous social media sites are not associated with well-defined user identities or profiles. In this study, our goals are two-fold: (i) to understand the nature (sensitivity, types) of content posted on anonymous social media sites and (ii) to investigate the differences between content posted on anonymous and non-anonymous social media sites like Twitter. To this end, we gather and analyze extensive content traces from Whisper (anonymous) and Twitter (non-anonymous) social media sites. We introduce the notion of *anonymity sensitivity* of a social media post, which captures the extent to which users think the post should be anonymous. We also propose a human annotator based methodology to measure the same for Whisper and Twitter posts. Our analysis reveals that anonymity sensitivity of most whispers (unlike tweets) is not binary. Instead, most whispers exhibit

While anonymous online forums have been in existence since the early days of the Internet, in the past, such forums were often devoted to certain sensitive topics or issues. In addition, its user population was relatively small and limited to technically sophisticated users with specific concerns or requirements to be anonymous. On the other hand, anonymous social media sites like Whisper[1] and Secret[2] provide a generic and easy-to-use platform for lay users to post their thoughts in relative anonymity. Thus, the advent and rapidly growing adoption of these sites provide us with an opportunity for the first time to investigate how large user populations make use of an anonymous public platform to post content.

In this paper, our goal is to better understand the characteristics of *content* posted on anonymous social media sites. Specifically, we introduce the notion of *anonymity sensitivity* to measure the sensitivity of content posted on such sites. Intuitively, anonymity sensitivity of a message cap-

Anonymity sensitivity

- Older users are more anonymity sensitive than younger ones.
- College educated users are more (statistically significant) anonymity sensitive than non-college educated users.
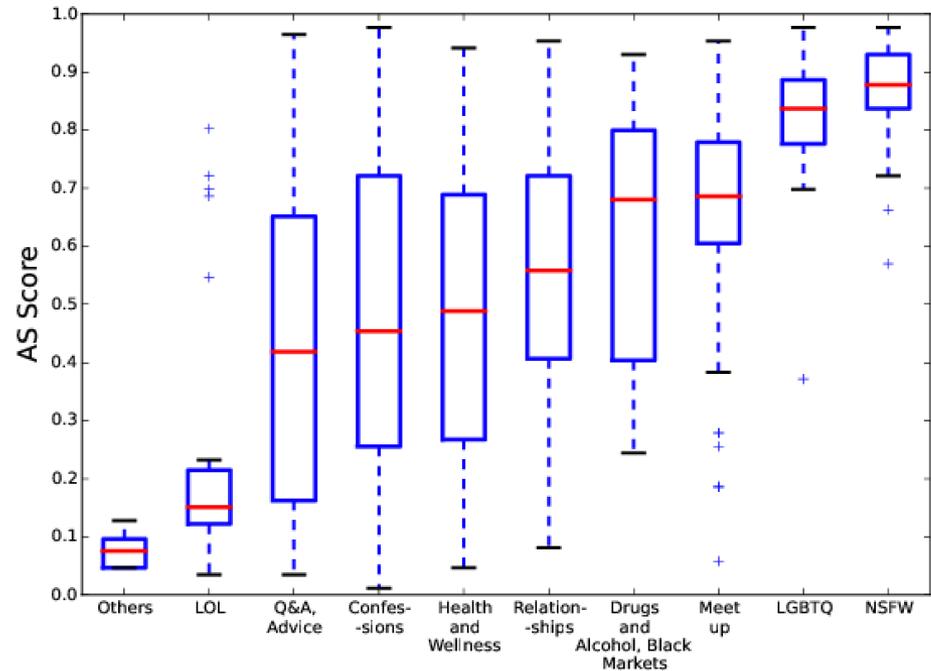- No gender difference.
  - WHY?

Figure 4: shows the distribution of *Anonymity Sensitivity Score* for all categories via box-and-whisker plot. *NSFW*, *Relationships* and *LGBTQ* categories contains highly sensitive content while *LOL* has low content sensitivity.

Correa et al 2015

"This is a Throwaway Account":
Temporary Technical Identities and Perceptions of Anonymity in a Massive Online Community

Alex Leavitt
Annenberg School for
Communication & Journalism,
University of Southern California
aleavitt@usc.edu

ABSTRACT

This paper explores temporary identities on social media platforms and individuals' uses of these identities with respect to their perceptions of anonymity. Given the research on multiple profile maintenance, little research has examined the role that some social media platforms play in affording users with temporary identities. Further, most of the research on anonymity stops short of the concept of varying perceptions of anonymity. This paper builds on these research areas by describing the phenomenon of temporary "throwaway accounts" and their uses on reddit.com, a popular social news site. In addition to ethnographic trace analysis to examine the contexts in which throwaway accounts are adopted, this paper presents a predictive model that suggests that perceptions of anonymity significantly shape the potential uses of throwaway accounts and that women are much more likely to adopt temporary identities than men.

features that afford a variety of boundary management techniques for information disclosure or identity performance. These affordances play a key role in how some users negotiate identity boundaries through the creation of multiple accounts, such as in moments when someone does not want particular information linked to the primary representation of him- or herself. I call these types of accounts "temporary technical identities," where the particular sociotechnical system affords an individual to create a provisional pseudonymous username. This paper explores one online platform, reddit.com, which permits users to create numerous accounts (under any pseudonym) easily, leading to the emergent phenomenon of "throwaway accounts."

This paper explores the typology and uses of temporary technical identities within the sociotechnical and cultural context of reddit guided by two research questions:

Does anonymity always have to be a design feature? Is it possible to be anonymous on a platform where this is not the case?

Over the last few years since the three studies, many new platforms have emerged. What are additional ways/cues people use to manage their identity?