Anonymity Effects: A Large-Scale Dataset from an Anonymous Social Media Platform

Mainack Mondal*
Indian Institute of Technology Kharagpur
Kharagpur, India
mainack@cse.iitkgp.ac.in

Denzil Correa*
Indraprastha Institute of Information Technology Delhi
Delhi, India
denzilc@iiitd.ac.in

Fabrício Benevenuto
Universidade Federal de Minas Gerais
Belo Horizonte, Brazil
fabricio@dcc.ufmg.br

ABSTRACT

Today online social media sites function as the medium of expression for billions of users. As a result, aside from conventional social media sites like Facebook and Twitter, platform designers introduced many alternative social media platforms (e.g., 4chan, Whisper, Snapchat, Mastodon) to serve specific userbases. Among these platforms, anonymous social media sites like Whisper and 4chan hold a special place for researchers. Unlike conventional social media sites, posts on anonymous social media sites are not associated with persistent user identities or profiles. Thus, these anonymous social media sites can provide an extremely interesting data-driven lens into the effects of anonymity on online user behavior. However, to the best of our knowledge, currently there are no publicly available datasets to facilitate research efforts on these *anonymity effects*.

To that end, in this paper, we aim to publicly release the first ever large-scale dataset from Whisper, a large anonymous online social media platform. Specifically, our dataset contains 89.8 Million Whisper posts (called "whispers") published between a 2-year period from June 6, 2014 to June 6, 2016 (when Whisper was quite popular). Each of these whispers contained both post text and associated metadata. The metadata contains information like coarse-grained location of upload and categories of whispers. We also present preliminary descriptive statistics to demonstrate a significant language and categorical diversity in our dataset. We leverage previous work as well as novel analysis to demonstrate that the whispers contain personal emotions and opinions (likely facilitated by a disinhibition complex due to anonymity). Consequently, we envision that our dataset will facilitate novel research ranging from understanding online aggression to detect depression within online populace.

CCS CONCEPTS

• **Information systems** → *Social networks.*

*The work was done when authors were part of MPI-SWS, Germany.

HT '20, July 13–15, 2020, Virtual Event, USA

ACM ISBN 978-1-4503-7098-1/20/07…$15.00
https://doi.org/10.1145/3372923.3404792

KEYWORDS

anonymity; social media; Whisper; pattern recognition; public dataset; hate speech

ACM Reference Format:
Mainack Mondal, Denzil Correa, and Fabrício Benevenuto. 2020. Anonymity Effects: A Large-Scale Dataset from an Anonymous Social Media Platform. In *Proceedings of the 31st ACM Conference on Hypertext and Social Media (HT '20), July 13–15, 2020, Virtual Event, USA.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3372923.3404792

1 INTRODUCTION

Online social media sites provide an accessible platform for billions of Internet users to contribute content and express their opinions. One outcome of such popularity is the enormous social-interaction dataset generated by the users of these platforms. This wealth of user-generated social data provides an excellent resource to study human behavior and group dynamics in research fields like computational sociology, psychology, data mining etc. [3, 15–17]. However, these interactions are inherently contextualized in a setting which enable interactions via platform provided features. One big component of such context is the social identities (i.e., the user profile in these platforms) which plays a key contributing role towards moderating social behavior. Studies have shown that social identities strongly influence the nature of participation and behavioral characteristics. Specifically, existing efforts in sociology [14, 18] point out that the existence of anonymous identities might expose hitherto unknown behaviors in a social setting—anonymity can increase disinhibition complex (users might be much less inhibited and express their otherwise suppressed feeling or ideas), but also can have no effects, or even decrease uninhibited behaviors.

Broadly, we can categorize social media sites into three types according to supported online social identities—non-anonymous (Facebook), pseudonymous (Twitter, Reddit) and anonymous (Whisper, Secret). These types of social identities can strongly influence the content and behavioral character of these sites. In this work, we focus on content posted by anonymous social identities. Recent work has shown that anonymous social media have fundamentally unique content and behavioral characteristics [5] compared to pseudonymous and non-anonymous sites—Content on anonymous social media posts have higher degree of sensitivity and distinct linguistic attributes. These earlier studies further found that, in terms of behavior, anonymous social media users post varying degrees of sensitive posts across different categories like confessions, relationships and spontaneous meetups. Such behaviors were absent in non-anonymous platforms. Other studies investigated additional characteristics like privacy guarantees of anonymous social media sites [1, 22]. In summary, these earlier studies demonstrate a unique and peculiar anonymity effect on social behavior and content.

Consequently, these studies point to the usefulness of these platforms to understand both desired (e.g., whistleblowing) and undesired effects (e.g., posting aggressive content to hurt a particular person of groups) of anonymity on user behavior. However, large-scale data-driven research on the effect of anonymous identities in a social setting is still not very widespread. To that end, in this ongoing work, we asked a simple yet important question—*How can we facilitate research on the effects of anonymity in a social setting?* We identify one problem with such research—lack of large-scale data. Specifically, in order to study characteristics of anonymous social communications, it would be pertinent to have access to a well curated dataset from an anonymous social media. Such a dataset would help accelerate research to understand how anonymous identities impact social dynamics. However, most such datasets are either simply not shared [1, 22] or limited to conversations from pseudonymous [6] and other non-anonymous sites [4, 13, 21].

Thus, in this work we release a public dataset of more than 89 million posts from an anonymous social media site viz. Whisper. The dataset was collected via continuously crawling the “latest” section of Whisper from June 2014 to June 2016, a time frame when Whisper claimed a soaring user base (10 million monthly active users in 2015) [10]. To the best of our knowledge, this is the first large dataset from an anonymous social media and we believe this dataset has the potential to provide a solid ground to understand the needs and pitfalls of anonymous communication.

In the rest of the paper, first we describe our data collection methodology and present a detailed description of the data elements in the collected data. Then we will present preliminary descriptive statistics from our dataset and describe our analysis as well as previous work demonstrated the uniqueness of this data. Finally, we will identify potential applications of the dataset and conclude.

2 COLLECTING DATA FROM AN ANONYMOUS SOCIAL MEDIA PLATFORM

We start this section by introducing Whisper—a popular anonymous platform. Then we will describe our strategy to collect Whisper posts over time and our approach to clean and anonymize the data.

2.1 Whisper: An anonymous online social media platform

Our aim in this work is to create a dataset to facilitate investigation of anonymity effects on social communications. Thus, we leveraged Whisper, a popular anonymous social media sharing site, as the platform of our focus. Whisper (launched in March 2012) is a mobile application in which users post anonymous messages called “whispers”. The popularity of Whisper increased over time with more than 2.5B page views by 2013, higher than even popular news websites like CNN [7] at that time. By 2013, Whisper acquired more than two million users and 45% of the users posted something every day [8]; Whisper reported a jump in acquiring new users in 2015 with 10 million monthly activity users in April 2015 [10]. News outlets identified whisper as “The place to go these days to vent, come clean, or peer into other people’s secrets” [9]. As of 2017 75% of Whisper users belong to the age group 18-34 and they are predominantly female [20].In this work, we focus on the formative years of Whisper (2014 to 2016) when it was gaining considerable popularity and acquiring a huge number of new users.

Although, Whisper users can only post messages via mobile phones, Whisper has a read-only web interface. In contrast to posts in traditional social media sites like Twitter, whispers do not contain identifiable user information. An initial username is randomly assigned by Whisper, but it is non-persistent i.e., users can change their usernames at any point of time. In addition, multiple users may choose to use the same username. Whisper users also do not have profile pages or related information and hence, a user cannot navigate whispers posted by any particular username. This anonymity aspect of whispers and the large number of whispers shared per day by the users, lend Whisper as a very attractive experimental testbed to investigate anonymous social communication.

Figure 1: shows an example of a posted whisper. It has 1,485 favorites, and was posted from North Carolina, USA.

Figure 1 shows example of a whisper. Each whisper is overlaid on an image which is randomly chosen or can be provided by the user. A user may also provide location information with whisper at different granularity levels. Each whisper can be “favorited” by another whisper user. We see that this particular whisper was originated from North Carolina, USA, favorited 1,485 times. Each whisper also contains a timestamp. Moreover, according to news reports, Whisper was moderated by both human and machine learning algorithms [12, 22]. Given the popularity of Whisper and its suitability as a platform to study anonymous social media behavior we decided to collect data from Whisper.

2.2 Data collection strategy

We collected Whisper data via the read-only web interface. This interface publicly releases the whispers. During the time of our data collection (starting from 2014) this web interface contained a section marked as “latest” which self-updated itself with the latest whispers as they were posted. In order to collect Whisper data, we decided to crawl this “latest” section continuously (no delay between two crawls) and accumulated these publicly available whispers.

We parsed the whisper data obtained from web interface, converted whispers as JSON formatted objects and stored those objects. We ran our data collection infrastructure for around two years starting from June 2014 and collected more than 89 million whispers.

2.3 Description of collected data

We stored the collected whisper objects in a flat file where each line of the file is one JSON object (corresponds to one unique whisper). Simply put, JSON objects are sets of key-value pairs; We decided to store the whispers as JSON objects due to ease-of-analysis and cross-language portability of JSON objects. In order to protect privacy of the Whisper users (although the data was publicly available), while creating shareable version of this dataset, we removed the unique alphanumeric identifiers provided by Whisper (more in Section 3).


```
{
  "places": [
    {"placetype": "region","name": "California"}
  ],
  "text": "It's hard to find a woman who likes dominant men
    but is also confident herself. Just something about being
    confident but choosing to give up control when I take it.
    ",
  "topics": ["being confident","dominant","find a woman"
    ,"confident","choosing","give up control","herself"],
  "ts": 1403287851758,
  "me2": 2,
  "serial": 1118315,
  "nickname": "Sir",
  "categories": [
    "Confessions"
  ]
}
```


Figure 2: shows an example of a whisper JSON object.

We present a whisper JSON object from our collected dataset in Figure 2 (this object corresponds to one line in our dataset). We describe the keys present in whisper JSON object in Table 1 which includes the text, whisper/user assigned categories, location metadata (places), #favorites, nickname, timestamp of the post and post’s unique serial number (assigned by us).

Note that, the keys presented on Table 1 is aggregated from our the whisper dataset containing 89,877,121 whispers. Thus not all whispers contain all keys. We also further anonymized (e.g., only keeping region and country level location, removing potential personally identifiable information) the data to for ethical considerations. We will elaborate this point more in Section 3. We further detect the language of whispers using a deep learning based language detection tool [2]. Whispers show a heavy skew towards English (possibly due to US-centric userbase).

.

We present the basic characteristics of our final whisper dataset is Table 2. We collected more than 89 million whispers over a two-year period when Whisper was hugely popular. We strongly believe that both the volume and the longitudinal nature of our dataset will help future studies to better understand anonymity effect on social communication over time. We request the reader to visit the following link for obtaining instructions to download our (anonymized) dataset of more than 89 million whispers:

https://github.com/Mainack/whisper-2014-2016-data-HT-2020

JSON object key name	Description	Data type
text	Text of the whisper	string
categories	List of categories assigned to this whisper by the users and/or Whisper’s internal algorithm	list
places	List of specific locations included the whisper object. Each location contains “placetype” and “name”. “placetype” can be either country (e.g., US) or region (e.g., US states)	list
me2	Number of favorites for the whisper	integer
nickname	Non-persistent anonymous username	string
ts	Unix timestamp in microseconds stating the upload time of the whisper	integer
serial	Unique serial number assigned by us. Replaces unique alpha-numeric id assigned by Whisper	integer
feeds	List of additional labels for the whisper text. Very few whispers contain this (not in Figure 2)	list

Table 1: Shows names, descriptions and data types of each key present in whisper objects.

Time Period of data collection	06/2014 – 06/2016
#whispers	89,877,121
# whispers written in English (%)	82,460,328 (91.7%)
# whispers with location (%)	52,465,255 (58.4%)
# whispers with assigned categories (%)	68,774,276 (76.5%)

Table 2: Basic statistics of our collected whispers.

3 ETHICAL CONSIDERATIONS

We recognize that the collection and sharing of publicly available data for research, although not uncommon, but can create some ethical concerns. However, investigating user communications in the presence of anonymity is an active research topic and this is the first large-scale data set which can substantially help those efforts (due to the scale, relevance and longitudinal nature of the data). However, these data can also pose a risk to the users in case it reveals their personally identifiable information (PII).

To that end, we took extra precautionary steps to actively minimize potential risks to users by further anonymization. We removed personally identifiable content like numeric ids assigned by Whisper, included images and urls to specific whisper posts. We also removed granular locations (like towns) and kept only country/region level locations. Finally, we used a regular expression-based approach [1] to identify potential PII like phone numbers, emails, ip addresses, bitcoin addresses, street addresses, zip codes, po boxes and ssn numbers in the whisper texts. Then we replaced all of those detected strings with easily-detectable placeholders like “[[POSSIBLE_BTC_ADDRESSES]]” in the whisper texts. We noted that only 0.5% whispers contained such possible PII (an upper bound

[1]We leveraged regular expressions mentioned in the codebase of https://github.com/mns-llc/bitsnarf

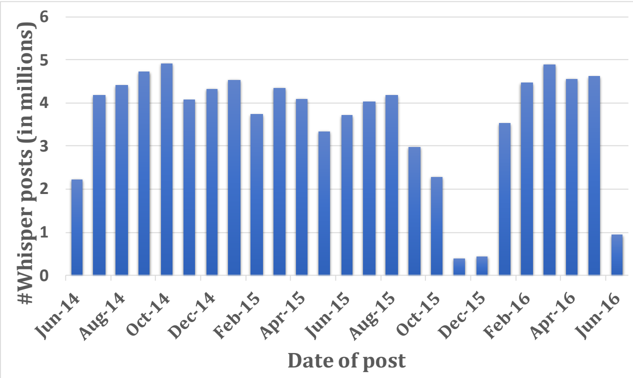


Figure 3: shows the number of whispers posted per month during our data collection period.

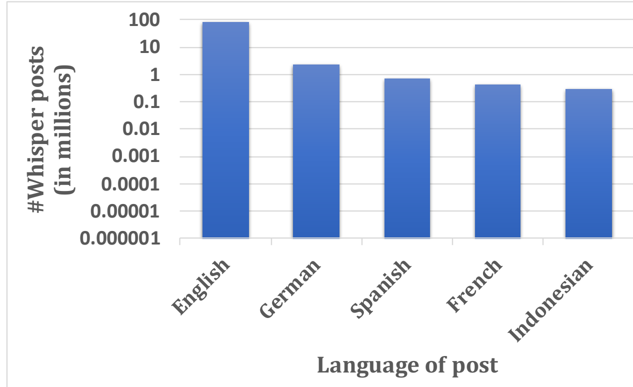


Figure 4: shows the number of whispers posted per language for five most frequent languages. The y-axis is in log-scale.

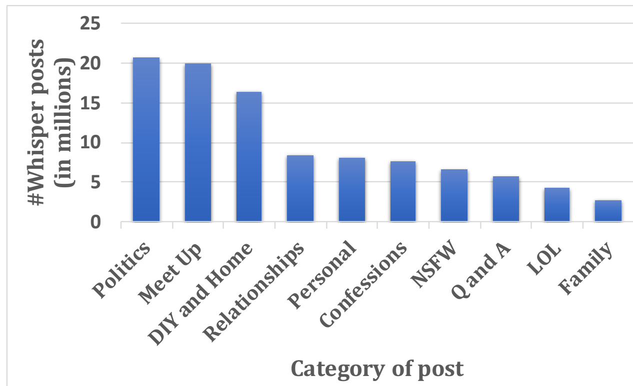


Figure 5: shows the number of whispers posted per category for ten most frequent categories.

since there might be false positives in PII detection). In summary, we thrived to address ethical considerations while releasing the first ever large-scale data from an anonymous social media site.

4 CHARACTERIZING WHISPER DATASET

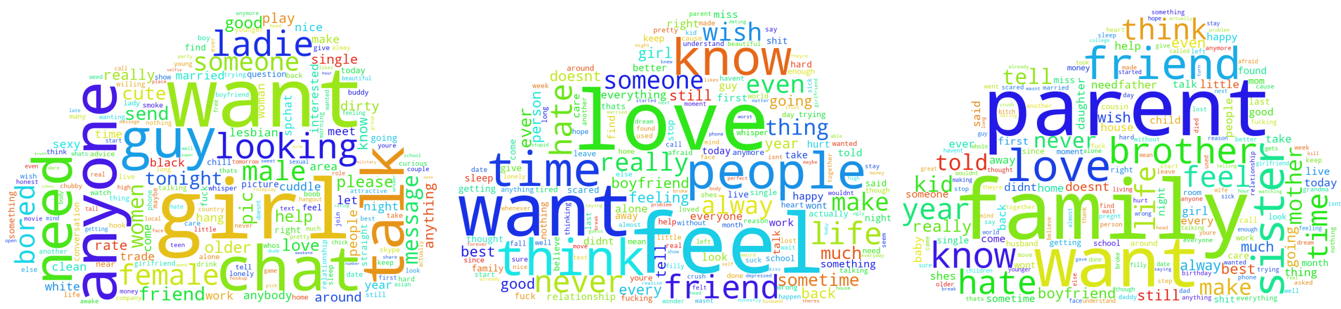
Now we will provide a brief characterization of our Whisper dataset. We will first check the volume of whispers posted over time and then demonstrate the diversity of our dataset in terms of language, most prominent categories and type of words used.

Number of whispers posted over time: We first check the number of whispers posted per month in our data collection period. We present the result in Figure 3 as a bar plot. We received on average 3.9 million whispers per month for the first 17 months (starting from June 2014). However, on November 2015 and December 2015, we received disproportionately low amount of whispers (average 0.4 million whispers per month). The reason is that, possibly due to storage issues during this period, the write times considerably increased, which in turn decrease the number of times we crawled "latest" section; we continued to receive similar volume of data as earlier from January 2016 onward once we changed our storage space. Furthermore, given our data collection start/end date did not align with start/end of month we received low volume of tweets on June 2014 and June 2016 (first and last month).

Language of whispers: Recall that majority (91.7%) of our whispers were written in English (shown in Table 2). However, a non-trivial number of whispers also written in non-English languages. To that end, we leveraged the output of a language detection tool [2] to identify five most frequent languages in our dataset. These languages and their corresponding number of whispers written in those languages is shown in Figure 4. We found that aside from English (82.5 million posts), whispers are also written in German (2.4 million posts), Spanish (0.7 million posts), French (0.4 million posts) and Indonesian (0.3 million posts). In total these five languages cover 96% whispers in our dataset. This result signifies the linguistic spread of whispers included in our dataset (in spite of the heavy skew towards English).

Categories of whispers: Each whisper object is labeled with (often multiple) Whisper/user-assigned categories (see Table 1). Next, we tried to check the ten most frequent categories assigned to whisper text in our dataset. These categories and their corresponding number of whispers is shown in Figure 5. Notably more than 15 million whispers are labeled with top two categories: "Politics" and "Meet up". This finding possibly hints at the fact that, users in Whisper post about their personal opinions and activities in their anonymous communications.

Next, we very briefly investigated the type of words included in different categories of whispers. Specifically, we randomly sample 100,000 whisper texts from each category, removed punctuation and stop words, and case-folded the text. Next, we create word clouds for each category using top 200 most frequent words. In Figure 6 6 we show three word clouds corresponding to the categories "Meet up", "Confessions" and "Family" (results for other labels are similar). We note that, these word clouds again hinted at more personal (and possibly sensitive) nature of the content like needs and desires of users. Earlier work also identifies similar patterns using LIWC and word trees [5]. In line to our finding they also note that whispers



(a) **Meet up** (b) **Confessions** (c) **Family**

Figure 6: shows the word clouds created with most frequent words for three categories of whispers.

are more personal, social, informal and sensitive than tweets across these categories. This further underlines the uniqueness of our data.

In summary, our dataset captures user activities over a long time and shows significant diversity in terms of language, categories and word usage. Next, we will conclude this paper by identifying potential applications of our dataset.

5 POTENTIAL APPLICATIONS

Our dataset contains an extensive amount of anonymous messages, exposing all kinds of user behavior. We hope this data might be useful for researchers on sociology and psychology to better comprehend human behavior under anonymity. Next, without any intention to be exhaustive, we describe a few potential directions that could be explored in this line.

Aggression and Offensive posts under Anonymity: As we mentioned in Section 1, anonymity can strongly influence user behavior by reducing societal boundaries in human attitudes. Existing research show that humans might turn aggressive and violent in situations in an environment that is less constrained by social norms and enable depersonalization of the self [19, 23]. Different systems have different content policies and approaches to moderate online aggression. The way in which digital platforms deal with content exclusion and moderation is usually complex as it may conflict with local legislation in each country, especially in terms of each country's parameters on freedom of expression. We hope our data can bring the perspective of aggression amplified by anonymity to the debate, fostering novel studies about online aggression and potentially influencing design of public policy for content moderation.

Disinhibition and Expressions of Depression It is expected that humans exhibit a higher disinhibition in communications under anonymity [14, 18]. A recent study about anonymity in Quora's question and answer website [11] unveil a few instances of such disinhibition, suggesting that users tend to openly express depression, anxiety, and very personal issues as they perceive they are anonymous. This is in line with what Whisper founders reported about their system, which even motivated them to start a separate nonprofit entity namely "Your Voice", dedicated for users to openly debate depression. We hope that researchers can leverage our dataset to uncover posts with signs of depression and other personal issues that users only comfortable to express anonymously.

Identifying Suicidal Tendencies. Although relatively less commonplace, but Whisper and other anonymous social media applications might be considered suitable forums for users to anonymously share their suicidal tendencies. In fact, there are anecdotal news stories pointing out a few instances of such cases[2]. We hope our large-scale whisper dataset will allow researchers to explore potential suicide tendencies expressed under the veil of anonymity and even develop machine learning models able to quick identify these kinds of messages to help the users.

6 CONCLUSION

To conclude, in this work we presented the data collection method for creating a novel large-scale dataset of millions of posts collected from an anonymous social media site. We demonstrate that our dataset is temporal, large-scale and contain significant diversity. We strongly believe that this Whisper dataset will be invaluable to the research community and open up novel research avenues. Specially we believe this data will be useful for future researchers and platform designers who thrive to shed light on and leverage the anonymity effects on user communications in a social setting.

REFERENCES

[1] Michael S Bernstein, Andrés Monroy-Hernández, Drew Harry, Paul André, Katrina Panovich, and Gregory G Vargas. 2011. 4chan and/b: An Analysis of Anonymity and Ephemerality in a Large Online Community. In *ICWSM*.

[2] Piotr Bojanowski, Edouard Grave, Armand Joulin, and Tomas Mikolov. 2016. Enriching Word Vectors with Subword Information. *arXiv preprint arXiv:1607.04606* (2016).

[3] Meeyoung Cha, Hamed Haddadi, Fabricio Benevenuto, and Krishna P. Gummadi. 2010. Measuring User Influence in Twitter: The Million Follower Fallacy. In *ICWSM*.

[4] Meeyoung Cha, Alan Mislove, and Krishna P. Gummadi. 2009. A Measurement-Driven Analysis of Information Propagation in the Flickr Social Network. In *WWW*.

[5] Denzil Correa, Leandro Silva, Mainack Mondal, Fabricio Benevenuto, and Krishna P. Gummadi. 2015. The Many Shades of Anonymity: Characterizing Anonymous Social Media Content. In *ICWSM*.

[6] Antigoni-Maria Founta, Constantinos Djouvas, Despoina Chatzakou, Ilias Leontiadis, Jeremy Blackburn, Gianluca Stringhini, Athena Vakali, Michael Sirivianos, and Nicolas Kourtellis. 2018. Large Scale Crowdsourcing and Characterization of Twitter Abusive Behavior. In *ICWSM*.

[2] https://www.businessinsider.com/suicide-notes-on-anonymous-apps-2014-5

[7] Liz Gannes. 2013. On Making Our Digital Lives More Real. http://allthingsd.com/20130802/im-so-over-oversharing-on-making-our-digital-lives-more-real/. (Accessed on April 2020).

[8] Erin Griffith. 2013. With 2 million users, "secrets app" Whisper launches on Android. http://pando.com/2013/05/16/with-2-million-users-secrets-app-whisper-launches-on-android/. (Accessed on April 2020).

[9] Kristine Johnson. 2017. Strangers Share Extreme Confessions On Whisper App. https://www.youtube.com/watch?v=-tYMNZUwuY4. (Accessed on April 2020).

[10] Matthew Lynley. 2015. After Passing 10 Million Monthly Active Users, Whisper Hires Its First President. https://techcrunch.com/2015/04/29/whisper-hires-its-first-president/#.hzgvrt:CTAX. (Accessed on April 2020).

[11] Binny Mathew, Ritam Dutt, Suman Kalyan Maity, Pawan Goyal, and Animesh Mukherjee. 2019. Deep Dive into Anonymity: Large Scale Analysis of Quora Questions. In *International Conference on Social Informatics*. Springer, 35–49.

[12] Harry Mccracken. 2016. Whisper's Master Of Content Moderation Is A Machine. https://www.fastcompany.com/3058148/whispers-master-of-content-moderation-is-a-machine. (Accessed on April 2020).

[13] Alan Mislove, Bimal Viswanath, Krishna P. Gummadi, and Peter Druschel. 2010. You Are Who You Know: Inferring User Profiles in Online Social Networks. In *WSDM*.

[14] Alain Pinsonneault and Nelson Heppel. 1997. Anonymity in group support systems research: A new conceptualization, measure, and contingency framework. *Journal of Management Information Systems* 14, 3 (1997), 89–108.

[15] K. Sassenberg. 2002. Common bond and common identity groups on the internet: Attachment and normative behavior in on-topic and off-topic chats. *Group Dynamics: Theory, Research, and Practice* 6, 1 (May 2002), 27–37.

[16] Leandro Araújo Silva, Mainack Mondal, Denzil Correa, Fabrício Benevenuto, and Ingmar Weber. 2016. Analyzing the Targets of Hate in Online Social Media. In *ICWSM*.

[17] Manya Sleeper, Rebecca Balebako, Sauvik Das, Amber Lynn McConahy, Jason Wiese, and Lorrie Faith Cranor. 2013. The Post that Wasn't: Exploring Self-Censorship on Facebook. In *CSCW*.

[18] John Suler. 2004. The online disinhibition effect. *Cyberpsychology & behavior* 7, 3 (2004), 321–326.

[19] John C. Turner. 1985. Social categorization and the self-concept: A social cognitive theory of group behavior. *Advances in Group Process* 2 (1985), 77–122.

[20] Neal Ungerleider. 2017. How Whisper Survives As Other Anonymous Social Apps Like Yik Yak Fail. https://www.fastcompany.com/40424834/how-whisper-survives-as-other-anonymous-social-apps-like-yik-yak-fail. (Accessed on April 2020).

[21] Bimal Viswanath, Alan Mislove, Meeyoung Cha, and Krishna P. Gummadi. 2009. On the Evolution of User Interaction in Facebook. In *WOSN*.

[22] Gang Wang, Bolun Wang, Tianyi Wang, Ana Nika, Haitao Zheng, and Ben Y. Zhao. 2014. Whispers in the Dark: Analyzing an Anonymous Social Network. In *IMC*.

[23] Philip G Zimbardo. 1969. The human choice: Individuation, reason, and order versus deindividuation, impulse, and chaos. *Nebraska Symposium on Motivation* 17 (1969), 237–307.