First Monday
PEER-REVIEWED JOURNAL ON THE INTERNET

A study of self-disclosure during the Coronavirus pandemic

by Taylor Blose, Prasanna Umar, Anna Squicciarini, and Sarah Rajtmajer

Abstract
We study observed incidence of self-disclosure in a large set of tweets representing user-led English-language conversation about the Coronavirus pandemic. Using an unsupervised approach to detect voluntary disclosure of personal information, we provide early evidence that situational factors surrounding the Coronavirus pandemic may impact individuals' privacy calculus. Text analyses reveal topical shift toward supportiveness and support-seeking in self-disclosing conversation on Twitter. We run a comparable analysis of tweets from Hurricane Harvey to provide context for observed effects and suggest opportunities for further study.

Contents

1. Introduction

The last thirteen months have seen global restrictions on movement unprecedented in modern times. Curfews, quarantines, stay-at-home and shelter-in-place orders, shutdowns and lockdowns have been variably instituted as COVID-19 case counts rise and fall. At one point in early April 2020, 91 percent of the world's population was living under travel restrictions of some kind and nearly 50 percent under lockdown (Connor, 2020). Unsurprisingly, there has been an unprecedented surge in online activity during this period (James, 2020). Much of the increased traffic extends beyond typical Internet surfing and video streaming, as people find ways to leverage online resources to stay connected with one another, personally and professionally. Social media usage has seen a 61 percent increase as people turn to these platforms to support routine social activities (Holmes, 2020).

It is to be expected that the expanded breadth and depth of online activity will magnify privacy risks for individual users. Beyond simply spending more time online, the emergence of virtual play dates, happy hours and book clubs suggests that during this time of social distancing, people are looking for ways to stay close (*e.g.*, "apart but together" campaigns). This may be particularly the case for the many individuals facing heightened anxiety, stress and depression due to social isolation, grief, financial insecurity, and of course health-related fears of the virus itself (Pfefferbaum and North, 2020; Gao, *et al.*, 2020; Galea, *et al.*, 2020; Rajkumar, 2020; Cauberghe, *et al.*, 2021; Drouin, *et al.*, 2020).

The literature on social communication suggests that interpersonal connectedness and relationship development is fundamentally facilitated through iterative *self-disclosure* (Altman and Taylor, 1973), that is, intentionally revealing personal information such as feelings, thoughts, and experiences to others (Derlaga and Berg, 1987). In fact, there is a robust literature on (routine) self-disclosure in online social media outside the particular domain of crisis (Bak, *et al.*, 2014; Joinson and Paine, 2007; Nguyen, *et al.*, 2012; Houghton and Joinson, 2012; Attrill and Jalil, 2011; Willems, *et al.*, 2020; Luo and Hancock, 2020). Research indicates that as users engage in discussion online, they leverage self-disclosure as a way to enhance immediate social rewards (Hallam and Zanella, 2017), increase legitimacy and likeability (Bak, *et al.*, 2012), and derive social support (Tidwell and Walther, 2002). A majority of users report that these rewards are mediated by concerns about privacy (Smith, *et al.*, 2011). Individual privacy concerns typically evolve over time, tied to the day's events and the longer arc of shifting norms (Acquisti, *et al.*, 2015; Adjerid, *et al.*, 2018). Ultimately, decisions to self-disclose are inherently personal and contextual, influenced by platform affordances (Joinson and Paine, 2007), audience (Bazarova and Choi, 2014; Choi and Bazarova, 2015), discussion topic (Umar, *et al.*, 2019), and peer effects (Barak and Gluck-Ofri, 2007).

Less is known about the evolution of users' sharing practices during public health emergencies. The crisis informatics literature has highlighted the role of social media during acute events for crisis management, public participation, and backchannel communication (Palen, 2008; Goolsby, 2010; Alexander, 2014). We know that victims of Hurricane Harvey sought assistance through social media, in some cases revealing their full names and addresses online (Seetharaman and Wells, 2017). However, what we are witnessing in the case of the Coronavirus pandemic is distinct from previous crises in important ways. COVID-19 is a global, relatively protracted acute threat. Unlike natural disasters or military engagements, the pandemic has left communication infrastructure intact. Digital outlets have become lifelines.

Our work addresses the following research questions. (*RQ1*) We ask whether and to what extent we see an increase in self-disclosing behaviors on Twitter during this time, specifically in posts related to the COVID-19 pandemic. (*RQ2*) We ask whether any observed deviations in self-disclosing behaviors correspond with acute events during the pandemic. Furthermore, we seek (*RQ3*) to identify the primary set of topics associated with self-disclosure in pandemic-related conversations and explore the implications such disclosures may therefore have on user privacy. Finally, we question (*RQ4*) whether themes and sharing behavior changes are crisis-centric or exhibited consistently across crises.

We analyze instances of self-disclosure in a dataset of 53,557,975 tweets representing conversations on Coronavirus-related topics. We define self-disclosure as the voluntary behavior of sharing personal information online within self-authored posts. We operationalize this definition using an unsupervised method for identifying personal information shared by an author about themselves. The approach we take identifies both subjective and objective personal information, and our analyses do not distinguish between the two.

Our analysis reveals a steep increase in instances of self-disclosure, particularly related to users' emotional states and personal experiences of the crisis. We compare observed self-disclosure patterns during the Coronavirus pandemic to observed self-disclosure patterns during a different crisis, *i.e.*, Hurricane Harvey in late summer 2017. Although hurricanes are an annual expectation, the landfall duration and subsequent impact of Hurricane Harvey created a crisis throughout communities in the south central region of the United States. This overwhelmed traditional emergency response infrastructure and affected citizens took to social media to seek emergency assistance (Sebastian, *et al.*, 2017; Smith, *et al.*, 2018). The Harvey study provides some context for these unprecedented times, suggesting similarities and differences that better inform our observations during the current crisis and implications for the privacy community.

2. Related work

Our work is situated within the general literature on self-disclosure. Historically the purview of psychologists' early work focused on the role of verbal disclosure in interpersonal relationships and identifying factors that may facilitate or inhibit these behaviors (see Cozby [1973] for a review). Research in the space of online interactions has sought to understand the actualization of self-disclosure in digitally-mediated social communication. Studies suggest that disclosure behaviors in online environments may be meaningfully different than their off-line counterparts, *e.g.*, anonymity and lack of non-verbal cues afforded by social media may encourage greater disclosure of sensitive information (Forest and Wood, 2012; Joinson, 2001). Similar findings are reported in Ma, *et al.* (2018) where the authors explore the impact of content intimacy on self-disclosure. It is well established for face-to-face communication that people disclose less as content intimacy increases, but this effect seems to be weakened in online interactions.

Our study is motivated by recent work positioning online self-disclosure as strategic behavior targeting social connectedness, self-expression, relationship development, identity clarification and social control (Bazarova and Choi, 2014; Gibbs, *et al.*, 2006; Abramova, *et al.*, 2017; De Choudhury and De, 2014). Voluntary disclosure of personal information has been associated with improved well-being, meaningfully related to increased informational and emotional support (Huang, 2016). In tandem, insights from the privacy literature suggest that stress may dampen privacy concerns related to self-disclosure on social media (Zhang, 2017; Zhang and Fu, 2020), and that self-disclosure may moderate the relationship between stressful life events and mental health (Zhang, 2017; Johnshoy, *et al.*, 2020). In tandem, these findings suggest that users may be engaging in increased self-disclosure during this time and drive our primary research question. Our work is responsive to a recent call to further understanding of online self-disclosure during the COVID-19 crisis (Nabity-Grover, *et al.*, 2020), and builds on initial studies of self-disclosure in response to psychosocial effects of the pandemic (Saha, *et al.*, 2020; Zhen, *et al.*, 2021).

Despite the “upsides”, *i.e.*, socially adaptive motivations for disclosure, we know that self-disclosure can come at a cost; leaving users exposed to identity theft, cyber fraud and other crimes (Hasan, *et al.*, 2013), discrimination in job searches, credit and visa applications (McGregor, *et al.*, 2018), and harassment and bullying (Peluchette, *et al.*, 2015). Of course, not all shared personal information is of equal concern with respect to privacy risk. More sensitive (Liu and Terzi, 2010) or surprising (Chen, *et al.*, 2013) disclosures may more meaningfully affect privacy risks.

Early work by Acquisti and Gross (2006) suggested that social network users were neither fully aware nor responsive to privacy risks. Over time, studies have captured a shift toward increased privacy awareness (Johnson, *et al.*, 2012; Vitak and Kim, 2014), but there remains great variability in information sharing behaviors amongst individuals and across platforms (Zhao, *et al.*, 2016). It has been shown that culture plays a role in disclosure decisions (Zhao, *et al.*, 2012; Krasnova, *et al.*, 2012; Trepte, *et al.*, 2017), as does gender (Sun, *et al.*, 2015) and socioeconomic status (SES) (Marwick, *et al.*, 2017). Overarchingly, the cost-benefit analyses underlying an individual’s decision to share in the presence of privacy risk is postulated by social exchange theory (Emerson, 1976) and re-framed in the context of online social networks as the so-called privacy calculus (Krasnova, *et al.*, 2010; Dienlin and Metzger, 2016). Critically, work in a number of domains suggests that contextual and situational factors, *e.g.*, trust, anonymity, financial incentives, are embedded within the privacy calculus (Joinson and Paine, 2012; Hann, *et al.*, 2007; Li, *et al.*, 2010). Amongst these factors, emotion has also been suggested to play a meaningful role in privacy behaviors (Laufer and Wolfe, 1977; Berendt, *et al.*, 2005; Li, *et al.*, 2017). This finding is in keeping with the general theory of feeling-as-information (Petty, *et al.*, 2001), whereby emotions serve as information cues directly invoking adaptive behaviors (Lazarus, 1991). Few studies have tried to link emotions with self-disclosure in circumscribed settings (Zhang, 2017; Zhang and Fu, 2020).

This work is complementary to the body of research in crisis informatics, which typically focuses on the related but fundamentally distinct problem of mining self-disclosed information for the purposes of identifying and deploying assistance and relief to impacted individuals and communities (see Muniz-Rodriguez, *et al.* [2020] for review). Threads in this field include social network analysis for disaster management (Zelenkauskaite, *et al.*, 2012; Imran, *et al.*, 2013; Takahashi, *et al.*, 2015; Chew and Eysenbach, 2010) and perceptions of crisis, including pandemics, mined through social media (Szomszor, *et al.*, 2011; Valaskivi, *et al.*, 2019). A robust literature is also emerging on the spread of misinformation in crisis (Starbird, *et al.*, 2014; Huang, *et al.*, 2015; Bursztyn, *et al.*, 2020) and we note the critical relevance of this work in the COVID-19 crisis we study here.

Our work dovetails with the literature on detection and tagging of self-disclosure in text (*e.g.*, Bak, *et al.*, 2014; Caliskan, *et al.*, 2014; Wang, *et al.*, 2016; Vasalou, *et al.*, 2011; Chow, *et al.*, 2008; Choi, *et al.*, 2013). Chow, *et al.* (2008) developed an association rules-based inference model that identified sensitive keywords which could be used to infer a private topic. Similarly, multiple studies utilized pattern or rule based methods to detect specific types of disclosures (Umar, *et al.*, 2019; Vasalou, *et al.*, 2011). Choi, *et al.* (2013) detected exposures of sensitive information such as location by looking for specific sentence patterns with occurrences of words like fly to, live in ,etc. used in conjunction with place names. Likewise, a study by Vasalou, *et al.* (2011) developed a privacy dictionary that differentiated between private and non private text. The detection was based on occurrence frequency of specific utterances associated with private categories. A recent method (Akiti, *et al.*, 2020) considers semantic role labeling to identify the lexical units and their semantic roles that signal self-disclosure.

Past work has attempted to classify self-disclosure by levels, or degree of disclosure. Caliskan, *et al.* (2014) used AdaBoost with a Naive Bayes classifier to detect privacy scores for Twitter users’ timelines. Bak, *et al.* (2014) applied modified Latent Dirichlet Allocation (LDA) topic models for semi-supervised classification of Twitter conversations into three self-disclosure levels: general, medium and high. Wang, *et al.* (2016) used regression models with extensive feature sets to detect degree of self-disclosure. Because the notion of sensitive information is based on user perception and context, studies on detection of self-disclosure levels are often difficult to generalize beyond their original context.

3. Datasets

Our primary dataset is a repository of tweet IDs corresponding to content posted on Twitter related to the Coronavirus pandemic (Chen, *et al.*, 2020). We collect a secondary dataset representing tweets during Hurricane Harvey to support comparative analyses and discussion of crisis-specific vs. crisis-general observations (*RQ4*).

3.1. COVID-19

Our primary dataset contains 508,088,777 tweet IDs for the period of activity from 21 January 2020 through 28 August 2020. Chen, *et al.* (2020) compiled tweets utilizing a combination of Twitter's Search API (for activity 21 January through 28 January) and Twitter's Streaming API (for activity 28 January through 31 July). The repository represents topically relevant tweets across the platform, canvassed based on designated keywords, as well as the full activity of selected accounts (See Tables 1, 2). Around 6 June 2020, the repository collection infrastructure transitioned to Amazon Web Services which generated a significant increase in tweet ID volume. No search parameters were adjusted or data gaps presented because of the transition; therefore, we analyzed the entirety of the dataset in a consistent manner. In analyses that follow, we focus our scope to highlight an important transitional period in the pandemic for individuals living in the United States by considering two sub-periods — prior to and after 11 March, the date of the World Health Organization's pandemic declaration and just two days preceding U.S. President Trump's declared national emergency.

Table 1: Keywords followed, by start date.

Keywords followed	Start date
Coronavirus, Koronavirus, Corona, CDC, Wuhancoronavirus, Wuhanlockdown, Ncov, Wuhan, N95, Kungflu, Epidemic, outbreak, Sinophobia, China	28 January 2020
covid-19	16 February 2020
corona virus	2 March 2020
covid, covid19, sars-cov-2	6 March 2020
COVID-19	8 March 2020
COVD, pandemic	12 March 2020
coronapocalypse, canceleverything, Coronials, SocialDistancingNow, Social Distancing, SocialDistancing	13 March 2020
panicbuy, panic buy, panicbuying, panic buying, 14DayQuarantine, DuringMy14DayQuarantine, panic shop, panic shopping, panicshop, InMyQuarantineSurvivalKit, panic-buy, panic-shop	14 March 2020
coronakindness	15 March 2020

quarantinelife, chinese virus, chinesevirus, stayhomechallenge, stay home challenge, sflockdown, DontBeASpreader, lockdown, lock down	16 March 2020
shelteringinplace, sheltering in place, staysafestayhome, stay safe stay home, trumppandemic, trump pandemic, flattenthecurve, flatten the curve, china virus, chinavirus	18 March 2020
quarentinelife, PPEshortage, saferathome, stayathome, stay at home, stay home, stayhome	19 March 2020
GetMePPE	21 March 2020
covidiot	26 March 2020
epitwitter	28 March 2020
pandemie	31 March 2020
wear a mask, wearamas, kung flu, covididiot	28 June 2020
COVID__19	9 July 2020

Table 2: Accounts followed, by start date.

Accounts followed	Start date
PneumoniaWuhan, CoronaVirusInfo,V2019N, CDCemergency, CDCgov, WHO, HHSGov, NIAIDNews	28 January 2020
drtedros	15 March 2020

Text and metadata corresponding to these 508,088,777 tweet IDs were obtained through rehydration using the Twarc [1] Python library. Of the IDs passed for rehydration, 461,259,923 were successfully rehydrated. The 9.21 percent loss represents deleted content, therefore irretrievable through Twitter's API.

For the purpose of measuring and studying self-disclosure, we filtered the corpus to capture tweets that represent original content posted by individual users. Specifically, we removed quoted tweets, retweets [2], as well as all tweets associated with verified accounts and the specific organizational accounts listed in Table 2. We narrowed our analysis to English-language content in order to reduce situational heterogenity and maintain confidence in our labeling approach, which has been developed and validated on English-language text. The resulting corpus, which forms the basis of our analyses, consists of 53,557,975 unique tweets.

3.2. Hurricane Harvey

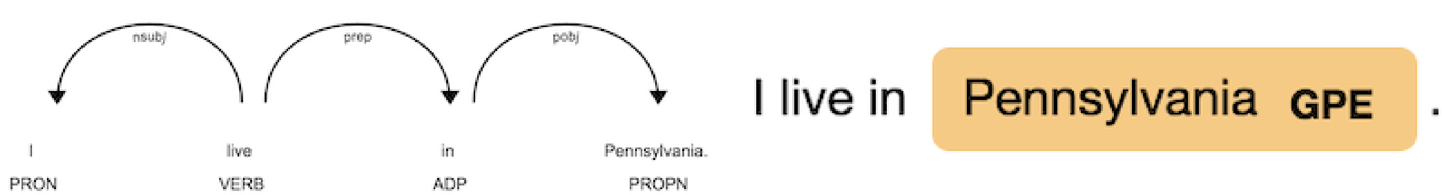
Our comparative dataset is a collection of 6,732,546 tweet IDs representing posted content inclusive of keywords “Hurricane Harvey”, “Harvey”, and/or “HurricaneHarvey” during a 12-day period after Harvey first made landfall, 25 August 2017, through 5 September 2017 (Alam, *et al.*, 2018). Similar to the COVID-19 dataset we hydrated the set of tweets using the Hydrator Tweet Retrieval Tool (v. 2.0) [3].

We experienced a 33 percent loss during data reconstitution (compare to 9.21 percent loss in the Coronavirus-related dataset), attributable to tweet and account deletion during the nearly three years which have passed since initial collection. We filter the resulting 4,379,462 tweets for original content in the same fashion as we handled the COVID-19 data — removing all quoted tweets, retweets, content from verified accounts, and non-English content as identified by Twitter. We pre-process the resulting 551,061 tweets and apply unsupervised labeling to detect instances of self-disclosure and topic analysis techniques as detailed in following Sections 4 and 5.

4. Automated detection of self-disclosure

4.1. Method

We use an unsupervised method (Umar, *et al.*, 2019) to detect instances of self-disclosure in our dataset. Consistent with the literature on detection of self-disclosure in text (see, *e.g.*, Bak, *et al.*, [2014]; Houghton and Joinson, [2012]; De Choudhury and De, [2014]; Wang, *et al.*, [2016]), we consider the presence of first-person pronouns. Specifically, we consider sentences containing self-reference as the subject, a category-related verb and associated named entity. Consider the example of location self-disclosure shown in Figure 1. A first person pronoun “I” is the self-referent subject of the sentence. It is used with a location-related verb “live” in the vicinity of the associated location entity “Pennsylvania”. Notably, subjective categories of self-disclosure such as interests and feelings do not have associated named entities. These are differentiated through rule-based schemas based on subject-verb pairs. The approach described is implemented in three phases: 1) subject, verb and object triplet extraction with awareness to voice (active or passive) in the sentence; 2) named entity recognition; 3) rule-based matching to established dictionaries. Our dictionaries are adopted from Umar, *et al.* (2019).



(a) Dependency tree

(b) Named entity recognition

Figure 1: Phases in self-disclosure categorization scheme [Umar, *et al.* (2019)].

As the proposed approach is based on sentence structure and syntactic resources (subject, verb, object and entities), it can be applied to any textual content. However, we acknowledge that tweets present unique characteristics. Due to character limits and consequent emerging norms of the platform, users more frequently engage acronyms and abbreviations (Han and Baldwin, 2011). Relatedly, sentence structure and syntax are noisier when compared to more verbose platforms (Boot, *et al.*, 2019). User mentions, hashtags, and graphic symbols are embedded within text. Considering these differences, we pre-process tweets as follows. All Unicode encoding errors are corrected. We remove markers associated with retweets (*e.g.*, “RT”) and filter user mentions and hashtags. Additionally, symbols like “&” and “$” are replaced with their respective word representations. E-mail addresses and phone numbers are replaced with placeholders “emailid” and “phonenumber”, while URLs are filtered. We also replace contractions in the tweets like “I’m” to “I am” and correct incorrect use of spacing between words. These pre-processing steps enable cleaner input to the detection algorithm.

4.2. Evaluation

While unsupervised classification is imperfect, our approach is both appropriate and effective for the task of labeling massive datasets like ours (see unsupervised approaches for similar tasks, *e.g.*, detecting offensive [Wiedemann, *et al.*, 2019], identifying ideology [Himelboim, *et al.*, 2013] and political party affiliation [Castro, *et al.*, 2017], and labeling stance on controversial topics [Darwish, *et al.*, 2020; Stefanov, *et al.*, 2020] on large Twitter datasets). Work developing supervised approaches for tagging self-disclosure in short text is ongoing, and all state of the art approaches still suffer from important limitations (see Related work).

To validate the unsupervised labeling approach in this context, we tested the method on a manually annotated subset of 5,000 tweets from our COVID-19 Twitter dataset. Each tweet was labelled by three raters on Amazon Mechanical Turk. Our labelling survey asked raters to evaluate different aspects of self-disclosure construct (Wang, *et al.*, 2016) on a scale from 1 (not at all) to 7 (completely). Specifically, raters were asked to rate the extent to which a tweet involved: 1) personal information; 2) personal thoughts; 3) personal feelings; 4) personal relationships; and, 5) the intimacy of the tweet (see Appendix A, Table A1). Intraclass correlation coefficient (ICC) showed fair to good (Cicchetti, 1994) agreement among raters for the first four labels (Good: Information (0.744), Feelings (0.665), Relations (0.672); Fair: Thoughts (0.446)). While, ICC for Intimacy was low, reflecting the particularly subjective nature of the label.

As our research questions center around incidence of self-disclosure, we consider four out of five dimensions (*i.e.*, excluding intimacy) of the self-disclosure construct that represent the categories of personal revelations. We binarized each of the self-disclosure categories namely, information, feelings, relations, and thoughts. A rating of “1” (Not at all) was taken as representative of no self-disclosure, while ratings greater than “1” represented presence of self-disclosure. Binary labels for each question were obtained for individual tweets through majority voting among three raters. A tweet was labelled self-disclosing if self-disclosure was noted for at least one of the four questions. Of 5,000 tweets, 4,297 were considered self-disclosing by this metric and the remaining 703 were not. We compared the performance of the automated approach to our manual labels. We obtained precision, recall and F1 scores of 93.3 percent, 62.9 percent, and 75.12 percent respectively. This performance provides support for the appropriation of the unsupervised labeling method in this work. We note that our research question centers around change over time rather than absolute quantification of self-disclosure, so that barring any temporal bias in the data and/or labelling approach, measures of change should be faithful.

5. Topic modeling

We consider topic modeling as a follow-up analysis to better understand topics of conversation in self-disclosing and non-self-disclosing tweets, and their evolution over time to address *RQ3*. We use Latent Dirichlet Allocation (LDA) (Blei, *et al.*, 2003), wherein each document (in our case, a single Tweet) in the corpus (set of tweets) is considered to be generated as a mixture of latent topics, where each topic is a distribution over words. For a document, each word is assigned topics according to a Dirichlet distribution. Iteratively cycling through each word in each document and all documents in the corpus, topic assignments are updated based on the prevalence of words across topics and the prevalence of topics in the document. Based on this process, topical coherence is assessed and final topic distributions for documents and word distributions for topics are generated.

In addition to the data cleaning steps described in Section 4, we pre-process tweets using tokenization, conversion to lower case, lemmatization and removal of punctuation and stopwords. Top six topics are generated from the resulting corpora using the LDA model within the Gensim Python library [4]. Gensim LDA is an unsupervised method to discover topical similarities within a corpus of text through analysis of semantic structure of the provided text.

We consider topic analyses over subsets of interest within the complete Coronavirus dataset. Namely, we explore unique topic models for tweets in two time windows pre- and post-March 11, 2020 (21 January — 11 March; 12 March — 15 May) and compare topical themes between tweets with presence or absence of detected instances of self-disclosure over the entirety of this subset. To better understand disclosure trends presented post-March 11, we also generate topic models over one-month windows for the remainder of the Coronavirus dataset: 16 May through 15 June, 16 June through 15 July, and 16 July through 15 August.

Through pyLDAvis (Sievert and Shirley, 2014) we generate intertopic distance maps (Appendix B) and top keywords for each generated topic. The intertopic distance maps visualize the topics in two-dimensional space with the area of the circles proportional to the amount of words belonging that topic across the dictionary. Overlapping topical circles show crossover in keywords identified, thus overlapping theme content. Distinct themes are evidenced by non-overlapping clusters spread out across the space. Key terms are classified through saliency and relevance. Saliency is a measure of how useful the term is for interpreting the topic while relevance is the measure at which a word belongs to a topic at the exclusion of being included in another topic. The relevance parameter lambda, λ, ranges from 0 to 1. A λ value closer to 0 identifies words more exclusive to the topic, while λ closer to 1 identifies terms more frequently presented in the topic, but also present in other topics. We use a λ value of 0.6, as this has been deemed optimal for interpreting topics (Sievert and Shirley, 2014).

6. Findings

Here we explore the frequency of self-disclosure tweets within the COVID-19 dataset, as well as topical themes for previously mentioned subsetted windows of time. We also present work on the Hurricane Harvey dataset for comparison analysis.

6.1. Analyzing COVID-19 disclosures

Of the 53,557,975 Coronavirus-related tweets analyzed, approximately 19.07 percent (10,215,752 tweets) contain elements of self-disclosure. Looking more closely at daily variance from January until June, we identify a significant transition point in activity around 11 March 2020, as shown in Figure 2. To fully answer *RQ1*, we confirm the significant behavioral change as a sustained pattern by overlaying 7-day and 30-day simple moving averages which smooth day-to-day variance by taking the average disclosure percentage value over the given time window.

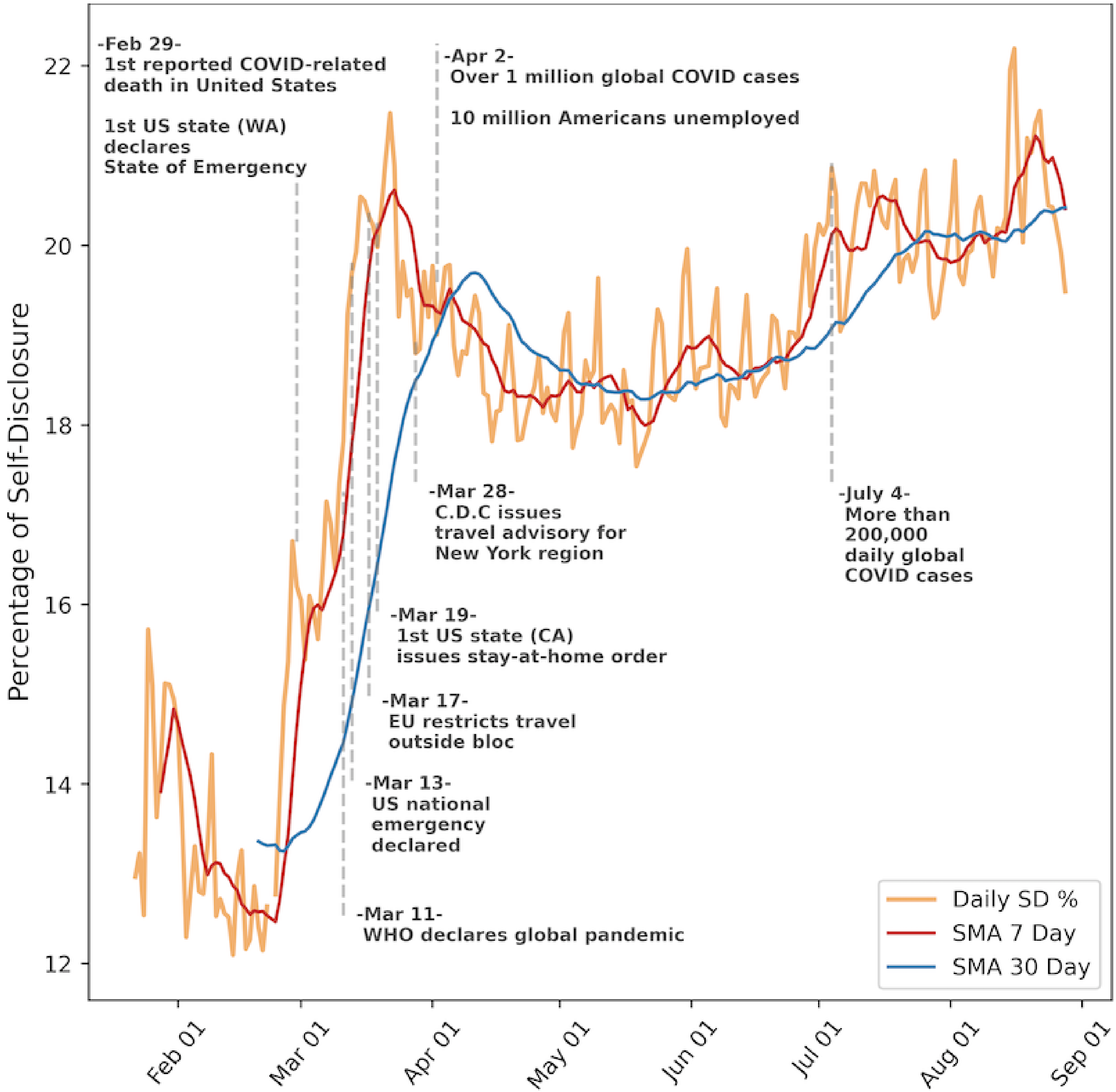


Figure 2: Percentage of tweets containing self-disclosure, assessed daily with 7- and 30-day simple moving averages.

We see disclosure change from the period 21 January through 11 March, where the average daily percentage of self-disclosing tweets is 14.63 peercent; compared to 12 March through 15 May, where the average daily percentage of self-disclosing tweets was 18.89 percent. Addressing *RQ2*, this change in activity coincides with an escalation in severity and increased global awareness of the crisis, with the World Health Organization (WHO) officially classifying Coronavirus as a pandemic (11 March). Current events coincident with observed changes in the rate of self-disclosure are noted on 13 March and 19 March when the United States officially declared a national state of emergency, and when the governor of California issued the first statewide 'stay-at-home' order, respectively. Self-disclosure activity remains high for the remainder of the dataset with an average daily percentage of 19.79 percent from 16 May through 28 August. As shown in Figure 2, additional inflection points are present coincident with acute events throughout the pandemic.

These observations suggest that situational context, in particular during crisis, may meaningfully influence short- and long-term disclosure behaviors.

We explore messaging around the Coronavirus pandemic through topic modeling, as outlined in Section 5. As discussed, self-disclosing behaviors increase steeply around 11 March, and interesting distinctions are noticeable in the topical breakdown comparing the periods just before (21 January 21 — 11 March) and after (12 March — 15 May) this date of interest, as reported in Table 3.

Table 3: Topical comparison of self-disclosing COVID-19 tweets before and after 11 March 2021.

Topic	Top six keywords
1	cdc,death,flu,test,number,report
2	get,like,fuck,shit,mask,die
3	trump,president,america,blame,economy,country
4	hand,buy,stock,wash,market,sell
5	hope,help,pray,god,epidemic,fight
6	travel,case,flight,quarantine,wuhan,confirm
(a) 21 January — 11 March 2020	
Topic	**Top six keywords**
1	distance,social,love,day,time,watch
2	corona,death,covid,virus,die,get
3	china,trump,country,virus,president,world
4	covid19,health,test,need,community,help
5	home,stay,mask,safe,order,wear
6	help,pay,sir,money,job,request
(b) 12 March — 15 May 2020	

Diving deeper to answer the presence of topical themes asked in *RQ3*, we examine that leading up to 11 March (Table 3a), self-disclosing conversation focused on general information (Topic 1) and sentimental impacts (Topic 2) representing 20 percent and 24 percent of all tweets, respectively. After 11 March (Table 3b), terms related to global and sentimental aspects of the crisis remain present (Topic 3, Topic 2) but prominent conversations also shift to managing the spread of the virus (Topic 5, Topic 1) and discussions about needs, help, thanks and support (Topic 4, Topic 6). In fact, Topics 4 and 6 together make up nearly half of the Tweets in that period, representing 23 percent and 24 percent of activity, respectively. This represents a shift from outward-looking to self-centric messaging as well as early evidence of emotional support and support seeking through disclosure.

Centered on the mid-March increase in disclosure behaviors, we also compare topical variance between disclosing and non-disclosing Coronavirus-related tweets for entirety of the aforementioned subsetted time window (21 January — 15 May; see Table 4). Generally, extracted topics reflect terminology pervasive in mainstream media at the onset of the crisis, including but not limited to expected impacts of COVID-19 (economy, market, lockdown, quarantine), recommendations (stay home, social distance, mask), and political impact (Trump, administration, democrat, government). Topic 1, the most prominent topic within the self-disclosing tweets, is represented by emotive terms (like, fuck, feel) not prominent in any non-self-disclosing themes.

Table 4: Topical comparison of self-disclosing and non-self-disclosing COVID-19 tweets through 15 May 2020.

Topic	Top six keywords
1	get,like,lockdown,day,fuck,feel
2	country,world,market,economy,chinese,government
3	test,case,death,flu,number,infect
4	mask,help,covid19,work,sir,support
5	trump,die,president,say,lie,know
6	home,stay,social,distance,safe,lockdown
(a) Self-disclosing tweets	
Topic	**Top six keywords**
1	virus,people,corona,like,time,spread
2	covid19,outbreak,health,help,pandemic,social
3	lockdown,home,stay,thank,safe,late
4	test,news,wuhan,patient,quarantine,hospital
5	case,new,death,report,confirm,update
6	trump,president,house,white,democrat,response
(b) Non-self-disclosing tweets	

We also analyze monthly subsets within the remaining timeline of our dataset, 15 May through 15 August, where the proportion of self-disclosing conversations remains elevated, and in fact, continues to gradually increase over time. The observed sustained higher rate of self-disclosure through the summer is particularly interesting. We speculate that users may have recalibrated their own sharing practices during the pandemic, and that this recalibration may represent a longer-term effect. Consistent with preceding time windows were themes of support seeking and political opinion. We observe the introduction of new themes related to social movements and education, while, discussions surrounding daily life (working/staying home) become the most prominent topic representing 38 percent and 41 percent of all self-disclosing tweets in the final two months of the collection. This again suggests a possible transition from discussions of anxiety and need to potentially coping with lifestyle changes and establishing a new normal. A listing of topics in this period is provided in Appendix A, Table A2.

6.2. Comparing COVID to Hurricane Harvey

To better understand RQ4 and the level of uniqueness attributable to the pandemic, we compare findings from the COVID and Hurricane Harvey datasets. Through analysis of the 551,061 tweets in the Hurricane Harvey dataset, we observe an average nine percent self-disclosure rate (49,595 tweets) over the 12-day collection period — substantially lower than the 19.07 percent observed in the Coronavirus-related dataset. Several factors might account for the difference. The current pandemic has been marked by efforts to maintain social distance and, we have proposed, increased levels of disclosure may be related to relationship-building with online cohorts. But what we may also be seeing is a reflection of a general trend toward greater self-disclosure in online social media over the course of nearly three years separating the two events.

Table 5: Example of self-disclosing tweets for selected topics from Table 3.

Period	Topic keywords	Example of self-disclosing tweet

Before 11 March	travel, case, flight, quarantine, wuhan, confirm	Booking Number: [Removed] and [Removed] Guest Name: [Removed] Query/Concern: with the recent travel advisories and uplift in the no of cases of covid-19, I would like to request a refund of my flight to/fro bangkok.
	hand, buy, stock, wash, market, sell	I work in a hospital and I am around sick people daily. Am I worried about the coronavirus? Nope not really. I get more worried about hiv/aids when I go into Ryan White center. I still say I should buy stock in hand sanitizer lol
	cdc, death, flu, test, number, report	October 2019 — february 2020 CDC Flu Tracking: 18,000 46,000 estimated deaths with 310,000 to 560,000 hospitalizations. I caught the flu in November last year on a United/SFO trip to Oakland. Passed to my family in Indiana over Thanksgiving.
After 11 March	home, stay, mask, safe, order, wear	I started wearing a mask about a month ago. Been out about 5 times. No one really looks at me because others are wearing masks too. Was at Walmart yesterday and none of the workers had gloves OR masks on! I was shocked but I am in Las Vegas, NV where this is no stay at home order.
	covid 19, health, test, need, community, helpt	*I am 22 yrs old and I tested positive for COVID-19* Do not underestimate this virus. I have never been so sick in my life. I am lucky and grateful to be as healthy as I am, but not everyone else is. Thank you for helping me to get tested. Everyone, please be safe!
	distance, social, love, day, time, watch	I started social distancing to protect my patients on Saturday. Since then, approximately my entire town has come to my home to see me. I am loved. And also reverse social distanced.

With respect to topical focus, we see important parallels between the two crises. As illustrated in Table 6a, self-disclosing tweets reveal emotional messaging centered on seeking immediate spiritual, physical, and monetary support (Topics 1, 2, 4, 5, 6) with top terms including “red cross”, “raise money”, “donation”, “relief”, and “prayer”. While present, these themes are less prominent in the non-self-disclosing data. This finding across both datasets suggests that support-seeking during crisis might be a driver of self-disclosure and play a meaningful role in users’ sharing practices. Also mirroring the Coronavirus dataset, we observe one self-disclosing topic representing politically motivated conversation (Topic 3).

Table 6: Topical comparison of self-disclosing and non-self-disclosing tweets during Hurricane Harvey.

Topic	Top six keywords
1	affect,heart,prayer,thought,texas,safe
2	relief,donation,effort,fund,jake,paul
3	sign,day,trump,petition,hear,tornado
4	know,god,love,need,think,deal

5	money,raise,houston,add,click,playlist
6	red,cross,team,american,save,best

(a) Self-disclosing tweets

Topic	Top six keywords
1	think,know,get,bad,feel,good
2	relief,effort,donation,fund,day,gas
3	family,safe,stay,add,lose,save
4	people,prayer,affect,thought,team,texas
5	support,red,wake,cross,bring,american
6	money, sign, raise, click, jake, paul

(b) Non-self-disclosing tweets

While there is topical overlap between the two classes of Harvey-related tweets, non-disclosing tweets (Table 6b) present a focus on crisis-specific contextual information with relevant keywords “flood”, “Texas”, and “Houston” (Topic 4).

7. Discussion and conclusion

The most striking observation in the analyses we have described is evidence of heightened and sustained levels of self-disclosure during the ongoing Coronavirus pandemic, contextualized by observed disclosure during Hurricane Harvey. This global crisis is unprecedented in a number of ways, one being the scale and scope of human interaction through social media. Concerns about privacy have been at the center of discussion in popular press (see, *e.g.*, Servick [2020]; Cellan-Jones [2020]; Lin and Martin [2020]), but most of this conversation has been about privacy tradeoffs related to location surveillance and individual health monitoring in service to public health. Many appear willing to sacrifice some privacy in hopes of stemming the spread of the disease and helping to accelerate the return to normalcy; others are not. Scholarly work has begun to propose “privacy first” decentralized approaches for COVID-related tracking and notification (see, *e.g.*, PACT [2020]; COVID Watch [2020]; Dp^{3}t [2020]).

We aim, with this work, to engage the privacy community in the work of better understanding heightened self-disclosure behaviors emergent in the Coronavirus pandemic and in crisis more broadly. Through our topic analysis we know that, during Hurricane Harvey, individuals took to social media in search of immediate aid. Today, they are leaning on their online social communities for ongoing engagement and support. Our analyses suggest that the current pandemic and its effects may have a sustained impact self-disclosure behaviors. Exploratory topic analysis reveals increased self-disclosure around sensitive topics (*e.g.*, health, money, help, support). An increase in disclosures containing such sensitive information has meaningful implications for privacy risks, especially during a time when users are already vulnerable (Liu and Terzi, 2010; Chen, *et al.*, 2013).

Isolation, economic uncertainty, and health-related anxiety pose serious threat to mental health and well-being (Holmes, *et al.*, 2020; Academy of Medical Sciences U.K., 2020), but the potential manifestations of psychological impact in the domain of voluntary self-disclosure are unknown. The existing literature hints at the role of mood and emotion in the privacy calculus, but these relationships have not be well established.

An open question is whether this increase in sharing practices will become a new norm, or whether they fade over time and we will ultimately return to pre-COVID baselines. If the higher volume of self-disclosure sustains beyond this crisis, we might look to social theory for explanations. Threshold models of collective behavior (see, *e.g.*, Granovetter [1978]; Centola, *et al.* [2018]; Wiedermann, *et al.* [2020]) may offer insights. As the crisis continues to unfold, we suggest that additional work should aim to further develop, refine and test these hypotheses.

About the authors

Taylor Blose is a a Ph.D. student in the College of Information Sciences and Technology at Pennsylvania State University.
E-mail: thb5018 [at] psu [dot] edu

Prasanna Umar is a Ph.D. student in the College of Information Sciences and Technology at that Pennsylvania State University.
E-mail: pxu3 [at] psu [dot] edu

Anna Squicciarini is an associate professor in the College of Information Sciences and Technology at Pennsylvania State University.
E-mail: acs20 [at] psu [dot] edu

Sarah Rajtmajer is an assistant professor in the College of Information Sciences and Technology at Pennsylvania State University.
Corresponding author: smr48 [at] psu [dot] edu

Acknowledgements

This work was supported by NSF RAPID award #2027757.

Notes

1. https://github.com/DocNow/twarc.

2. Retweets are identified through the existence of the "retweeted_status" field in the tweet object returned by the Twitter API. Tweets beginning with the string 'RT @' were also treated as retweeted records.

3. https://github.com/DocNow/hydrator.

4. https://radimrehurek.com/gensim/models/ldamodel.html.

References

Olga Abramova, Amina Wagner, Hanna Krasnova, and Peter Buxmann, 2017. "Understanding self-disclosure on social networking sites — A literature review," *AMCIS 2017 Proceedings*, at https://aisel.aisnet.org/amcis2017/AdoptionIT/Presentations/30/, accessed 20 June 2021.

Academy of Medical Sciences U.K., 2020. "Survey results: Understanding people's concerns about the mental health impacts of the cOVID-19 pandemic," at https://acmedsci.ac.uk/file-download/99436893, accessed 20 June 2021.

Alessandro Acquisti and Ralph Gross, 2006. "Imagined communities: Awareness, information sharing, and privacy on the Facebook," In: George Danezis and Philippe Golle (editors). *Privacy enhancing technologies. Lecture Notes in Computer Science*, volume 4258. Berlin: Springer, pp. 36–58.
doi: https://doi.org/10.1007/11957454_3, accessed 20 June 2021.

Alessandro Acquisti, Laura Brandimarte, and George Loewenstein, 2015. "Privacy and human behavior in the age of information," *Science*, volume 347, number 6221 (30 January), pp. 509–514.
doi: https://doi.org/10.1126/science.aaa1465, accessed 20 June 2021.

Idris Adjerid, Eyal Peer, and Alessandro Acquisti, 2018. "Beyond the privacy paradox: Objective versus relative risk in privacy decision making," *MIS Quarterly*, volume 42, number 2, pp. 465–488.
doi: https://doi.org/10.25300/MISQ/2018/14316, accessed 20 June 2021.

Chandan Akiti, Anna Squicciarini, and Sarah Rajtmajer, 2020. "A semantics-based approach to disclosure classification in user-generated online content," *Findings of the Association for Computational Linguistics: EMNLP 2020*, pp. 3,490–3,499.
doi: https://doi.org/10.18653/v1/2020.findings-emnlp.312, accessed 20 June 2021.

Firoj Alam, Ferda Ofli, Muhammad Imran, and Michael Aupetit, 2018. "A Twitter tale of three hurricanes: Harvey, Irma, and Maria," *ISCRAM 2018 Conference Proceedings — 15th International Conference on Information Systems for Crisis Response and Management*, pp. 553–572.

David E. Alexander, 2014. "Social media in disaster risk reduction and crisis management," *Science and Engineering Ethics*, volume 20, number 3, pp. 717–733.
doi: https://doi.org/10.1007/s11948-013-9502-z, accessed 20 June 2021.

Irwin Altman and Dalmas A Taylor, 1973. *Social penetration: The development of interpersonal relationships*. New York: Holt, Rinehart and Winston.

Alison Attrill and Rahul Jalil, 2011. "Revealing only the superficial me: Exploring categorical self-disclosure online," *Computers in Human Behavior*, volume 27, number 5, pp. 1,634–1,642.
doi: https://doi.org/10.1016/j.chb.2011.02.001, accessed 20 June 2021.

Jin Yeong Bak, Chin-Yew Lin, and Alice Oh. "Self-disclosure topic model for classifying and analyzing Twitter conversations," *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 1,986–1,996.
doi: https://doi.org/10.3115/v1/D14-1213, accessed 20 June 2021.

Jin Yeong Bak, Suin Kim, and Alice Oh, 2012. "Self-disclosure and relationship strength in Twitter conversations," *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Short Papers*, volume 2, pp. 60–64, and at https://www.aclweb.org/anthology/P12-2012/, accessed 20 June 2021.

Azy Barak and Orit Gluck-Ofri, 2007. "Degree and reciprocity of self-disclosure in online forums," *CyberPsychology & Behavior*, volume 10, number 3, pp. 407–417.
doi: https://doi.org/10.1089/cpb.2006.9938, accessed 20 June 2021.

Natalya N. Bazarova and Yoon Hyung Choi, 2014. "Self-disclosure in social media: Extending the functional approach to disclosure motivations and characteristics on social network sites," *Journal of Communication*, volume 64, number 4, pp. 635–657.
doi: https://doi.org/10.1111/jcom.12106, accessed 20 June 2021.

Bettina Berendt, Oliver Günther, and Sarah Spiekermann, 2005. "Privacy in e-commerce: Stated preferences vs. actual behavior," *Communications of the ACM*, volume 48, number 4, pp. 101–106.
doi: https://doi.org/10.1145/1053291.1053295, accessed 20 June 2021.

David M. Blei, Andrew Y. Ng, and Michael I. Jordan, 2003. "Latent Dirichlet allocation," *Journal of Machine Learning Research*, volume 3, pp. 993–1,022, and at https://www.jmlr.org/papers/volume3/blei03a/blei03a.pdf, accessed 20 June 2021.

Arnout B. Boot, Erik Tjong Kim Sang, Katinka Dijkstra, and Rolf A. Zwaan, 2019. "How character limit affects language usage in tweets," *Palgrave Communications*, volume 5, article number 76.
doi: https://doi.org/10.1057/s41599-019-0280-3, accessed 20 June 2021.

Leonardo Bursztyn, Aakaash Rao, Christopher Roth, and David Yanagizawa-Drott, 2020. "Misinformation during a pandemic," *University of Chicago, Becker Friedman Institute for Economics, Working Paper* (1 September), at https://bfi.uchicago.edu/working-paper/2020-44/, accessed 20 June 2021.

Aylin Caliskan Islam, Jonathan Walsh, and Rachel Greenstadt, 2014. "Privacy detective: Detecting private information and collective privacy behavior in a large social network," *WPES '14: Proceedings of the 13th Workshop on Privacy in the Electronic Society*, pp. 35–46.
doi: https://doi.org/10.1145/2665943.2665958, accessed 20 June 2021.

Rodrigo Castro, Leonardo Kuffó, and Carmen Vaca, 2017. "Back to# 6D: Predicting Venezuelan states political election

results through Twitter," *2017 Fourth International Conference on eDemocracy & eGovernment (ICEDEG)*, pp. 148–153.
doi: https://doi.org/10.1109/ICEDEG.2017.7962525, accessed 20 June 2021.

Verolien Cauberghe, Ini Van Wesenbeeck, Steffi De Jans, Liselot Hudders, and Koen Ponnet, 2021. "How adolescents use social media to cope with feelings of loneliness and anxiety during COVID-19 lockdown," *Cyberpsychology, Behavior, and Social Networking*, volume 24, number 4, pp. 250–257.
doi: https://doi.org/10.1089/cyber.2020.0478, accessed 20 June 2021.

Rory Cellan-Jones, 2020. "Coronavirus: Privacy in a pandemic," *BBC News* (2 April), at https://www.bbc.com/news/technology-52135916, accessed 20 June 2021.

Damon Centola, Joshua Becker, Devon Brackbill, and Andrea Baronchelli, 2018. "Experimental evidence for tipping points in social convention," *Science*, volume 360, number 6393 (8 June), pp. 1,116–1,119.
doi: https://doi.org/10.1126/science.aas8827, accessed 20 June 2021.

Emily Chen, Kristina Lerman, and Emilio Ferrara, 2020. "Tracking social media discourse about the COVID-19 pandemic: Development of a public Coronavirus Twitter data set," *arXiv*:2003.07372 (2 June), at https://arxiv.org/abs/2003.07372, accessed 20 June 2021.

Terence Chen, Abdelberi Chaabane, Pierre Ugo Tournoux, Mohamed-Ali Kaafar, and Roksana Boreli, 2013. "How much is too much? leveraging ads audience estimation to evaluate public profile uniqueness," In: Emiliano De Cristofaro and Matthew Wright (editors). *Privacy enhancing technologies. Lecture Notes in Computer Science*, volume 7981. Berlin: Springer, pp. 225–244.
doi: https://doi.org/10.1007/978-3-642-39077-7_12, accessed 20 June 2021.

Cynthia Chew and Gunther Eysenbach, 2010. "Pandemics in the age of Twitter: Content analysis of tweets during the 2009 H1N1 outbreak," *PloS ONE*, volume 5, number 11, e14118.
doi: https://doi.org/10.1371/journal.pone.0014118, accessed 20 June 2021.

Dongjin Choi, Jeongin Kim, Xeufeng Piao, and Pankoo Kim, 2013. "Text analysis for monitoring personal information leakage on Twitter," *Journal of Universal Computer Science*, volume 19, number 16, pp. 2,472–2,485.
doi: https://doi.org/10.3217/jucs-019-16-2472, accessed 20 June 2021.

Yoon Hyung Choi and Natalya N. Bazarova, 2015. "Self-disclosure characteristics and motivations in social media: Extending the functional model to multiple social network sites," *Human Communication Research*, volume 41, number 4, pp. 480–500.
doi: https://doi.org/10.1111/hcre.12053, accessed 20 June 2021.

Richard Chow, Philippe Golle, and Jessica Staddon, 2008. "Detecting privacy leaks using corpus-based association rules," *KDD '08: Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 893–901.
doi: https://doi.org/10.1145/1401890.1401997, accessed 20 June 2021.

Domenic V. Cicchetti, 1994. "Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology," *Psychological Assessment*, volume 6, number 4, pp. 284–290.
doi: https://doi.org/10.1037/1040-3590.6.4.284, accessed 20 June 2021.

Phillip Connor, 2020. "More than nine-in-ten people worldwide live in countries with travel restrictions amid COVID-19," *Pew Research Center* (1 April), at https://policycommons.net/artifacts/616270/more-than-nine-in-ten-people-worldwide-live-in-countries-with-travel-restrictions-amid-covid-19/, accessed 20 June 2021.

COVID Watch, 2020. "COVID Watch," at https://policycommons.net/artifacts/616270/more-than-nine-in-ten-people-worldwide-live-in-countries-with-travel-restrictions-amid-covid-19/, accessed 20 June 2021.

Paul C. Cozby, 1973. "Self-disclosure: A literature review," *Psychological Bulletin*, volume 79, number 2, pp. 73–91.
doi: https://doi.org/10.1037/h0033950, accessed 20 June 2021.

Kareem Darwish, Peter Stefanov, Michaël Aupetit, and Preslav Nakov, 2020. "Unsupervised user stance detection on Twitter," *Proceedings of the International AAAI Conference on Web and Social Media*, volume 14, pp. 141–152, and at https://ojs.aaai.org//index.php/ICWSM/article/view/7286, accessed 20 June 2021.

Munmun De Choudhury and Sushovan De, 2014. “Mental health discourse on Reddit: Self-disclosure, social support, and anonymity,” *Eighth International AAAI Conference on Weblogs and Social Media*, volume 8, at https://ojs.aaai.org/index.php/ICWSM/article/view/14526, accessed 20 June 2021.

Tobias Dienlin and Miriam J. Metzger, 2016. “An extended privacy calculus model for SNSs: Analyzing self-disclosure and self-withdrawal in a representative U.S. sample,” *Journal of Computer-Mediated Communication*, volume 21, number 5, pp. 368–383.
doi: https://doi.org/10.1111/jcc4.12163, accessed 20 June 2021.

Valerian J. Derlaga and John H. Berg (editors), 1987. *Self-disclosure: Theory, research, and therapy*. New York: Springer Science + Business Media.
doi: https://doi.org/10.1007/978-1-4899-3523-6, accessed 20 June 2021.

Dp^{3}t, 2020. “Decentralized privacy-preserving proximity tracing,” at https://github.com/DP-3T/documents/blob/master/DP3T%20White%20Paper.pdf, accessed 20 June 2021.

Michelle Drouin, Brandon T McDaniel, Jessica Pater, and Tammy Toscos, 2020. “How parents and their children used social media and technology at the beginning of the COVID-19 pandemic and associations with anxiety,” *Cyberpsychology, Behavior, and Social Networking*, volume 23, number 11, pp. 727–736.
doi: https://doi.org/10.1089/cyber.2020.0284, accessed 20 June 2021.

Richard M. Emerson, 1976. “Social exchange theory,” *Annual Review of Sociology*, volume 2, pp. 335–362.
doi: https://doi.org/10.1146/annurev.so.02.080176.002003, accessed 20 June 2021.

Amanda L. Forest and Joanne V. Wood, 2012. “When social networking is not working: Individuals with low self-esteem recognize but do not reap the benefits of self-disclosure on Facebook,” *Psychological Science*, volume 23, number 3, pp. 295–302.
doi: https://doi.org/10.1177/0956797611429709, accessed 20 June 2021.

Sandro Galea, Raina M. Merchant, and Nicole Lurie, 2020. “The mental health consequences of COVID-19 and physical distancing: The need for prevention and early intervention,” *JAMA Internal Medicine*, volume 180, number 6, pp. 817–818.
doi: https://doi.org/10.1001/jamainternmed.2020.1562, accessed 20 June 2021.

Junling Gao, Pinpin Zheng, Yingnan Jia, Hao Chen, Yimeng Mao, Suhong Chen, Yi Wang, Hua Fu, and Junming Dai. 2020. “Mental health problems and social media exposure during COVID-19 outbreak,” *PLoS ONE*, volume 15, number 4, e0231924 (16 April).
doi: https://doi.org/10.1371/journal.pone.0231924, accessed 20 June 2021.

Jennifer L. Gibbs, Nicole B. Ellison, and Rebecca D. Heino, 2006. “Self-presentation in online personals: The role of anticipated future interaction, self-disclosure, and perceived success in Internet dating,” *Communication Research*, volume 33, number 2, pp. 152–177.
doi: https://doi.org/10.1177/0093650205285368, accessed 20 June 2021.

Rebecca Goolsby, 2010. “Social media as crisis platform: The future of community maps/crisis maps,” *ACM Transactions on Intelligent Systems and Technology*, volume 1, number 1, article number 7, pp. 1–11.
doi: https://doi.org/10.1145/1858948.1858955, accessed 20 June 2021.

Mark Granovetter, 1978. “Threshold models of collective behavior,” *American Journal of Sociology*, volume 83, number 6, pp. 1,420–1,443.
doi: https://doi.org/10.1086/226707, accessed 20 June 2021.

Cory Hallam and Gianluca Zanella, 2017. “Online self-disclosure: The privacy paradox explained as a temporally discounted balance between concerns and rewards,” *Computers in Human Behavior*, volume 68, pp. 217–227.
doi: https://doi.org/10.1016/j.chb.2016.11.033, accessed 20 June 2021.

Bo Han and Timothy Baldwin, 2011. “Lexical normalisation of short text messages: Makn sens a# Twitter,” *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, volume 1, pp. 368–378, and at https://www.aclweb.org/anthology/P11-1038/, accessed 20 June 2021.

Il-Horn Hann, Kai-Lung Hui, Sang-Yong Tom Lee, and Ivan P.L. Png, 2007. “Overcoming online information privacy concerns: An information-processing theory approach,” *Journal of Management Information Systems*, volume 24, number 2, pp. 13–42.
doi: https://doi.org/10.2753/MIS0742-1222240202, accessed 20 June 2021.

Omar Hasan, Benjamin Habegger, Lionel Brunie, Nadia Bennani, and Ernesto Damiani, 2013. “A discussion of privacy challenges in user profiling with big data techniques: The EEXCESS use case,” *2013 IEEE International Congress on Big Data*, pp. 25–30.
doi: https://doi.org/10.1109/BigData.Congress.2013.13, accessed 20 June 2021.

Itai Himelboim, Stephen McCreery, and Marc Smith, 2013. “Birds of a feather tweet together: Integrating network and content analyses to examine cross-ideology exposure on Twitter,” *Journal of Computer-Mediated Communication*, volume 18, number 2, pp. 154–174.
doi: https://doi.org/10.1111/jcc4.12001, accessed 20 June 2021.

Emily A. Holmes, Rory C. O’Connor, V. Hugh Perry, Irene Tracey, Simon Wessely, Louise Arseneault, Clive Ballard, Helen Christensen, Roxane Cohen Silver, Ian Everall, Tamsin Ford, Ann John, Thomas Kabir, Kate King, Ira Madan, Susan Michie, Andrew K. Przybylski, Roz Shafran, Angela Sweeney, Carol M. Worthman, Lucy Yardley, Katherine Cowan, Claire Cope, Matthew Hotopf, and Ed Bullmore, 2020. “Multidisciplinary research priorities for the COVID-19 pandemic: A call for action for mental health science,” *Lancet Psychiatry*, volume 7, number 6, pp. 547–560.
doi: https://doi.org/10.1016/S2215-0366(20)30168-1, accessed 20 June 2021.

Ryan Holmes, 2020. “Is COVID-19 social media’s levelling up moment?” *Forbes* (24 April), at https://www.forbes.com/sites/ryanholmes/2020/04/24/is-covid-19-social-medias-levelling-up-moment/, accessed 20 June 2021.

David J. Houghton and Adam N. Joinson, 2012. “Linguistic markers of secrets and sensitive self-disclosure in Twitter,” *2012 45th Hawaii International Conference on System Sciences*, pp. 3,480–3,489.
doi: https://doi.org/10.1109/HICSS.2012.415, accessed 20 June 2021.

Hsin-Yi Huang, 2016. “Examining the beneficial effects of individual’s self-disclosure on the social network site,” *Computers in Human Behavior*, volume 57, pp. 122–132.
doi: https://doi.org/10.1016/j.chb.2015.12.030, accessed 20 June 2021.

Y. Linlin Huang, Kate Starbird, Mania Orand, Stephanie A. Stanek, and Heather T. Pedersen, 2015. “Connected through crisis: Emotional proximity and the spread of misinformation online,” *CSCW ’15: Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, pp. 969–980.
doi: https://doi.org/10.1145/2675133.2675202, accessed 20 June 2021.

Muhammad Imran, Shady Elbassuoni, Carlos Castillo, Fernando Diaz, and Patrick Meier, 2013. “Extracting information nuggets from disaster-related messages in social media,” *Proceedings of the Tenth International ISCRAM Conference*, at https://mimran.me/papers/imran_shady_carlos_fernando_patrick_iscram2013.pdf, accessed 20 June 2021.

Meg James, 2020. “FCC oks Verizon request for more capacity as network use spikes,” *Government Technology* (19 March), at https://www.govtech.com/news/fcc-oks-verizon-request-for-more-capacity-as-network-use-spikes.html, accessed 20 June 2021.

Quinn Johnshoy, Erin Moroze, Isabella Kaser, Aleina Tanabe, Connor Adkisson, Samantha Hutzley, Celeste Cole, Sherlynn Garces, Klaudia Stewart, and Jay Campisi, 2020. “Social media use following exposure to an acute stressor facilitates recovery from the stress response,” *Physiology & Behavior*, volume 223, 113012 (1 September).
doi: https://doi.org/10.1016/j.physbeh.2020.113012, accessed 20 June 2021.

Maritza Johnson, Serge Egelman, and Steven M. Bellovin, 2012. “Facebook and privacy: It’s complicated,” *SOUPS ’12: Proceedings of the Eighth Symposium on Usable Privacy and Security*, article number 9, pp. 1–15.
doi: https://doi.org/10.1145/2335356.2335369, accessed 20 June 2021.

Adam N. Joinson, 2001. “Self-disclosure in computer-mediated communication: The role of self-awareness and visual anonymity,” *European Journal of Social Psychology*, volume 31, number 2, pp. 177–192.
doi: https://doi.org/10.1002/ejsp.36, accessed 20 June 2021.

Adam N. Joinson and Carina B. Paine, 2012. “Self-disclosure, privacy and the Internet,” In: Adam N. Joinson, Katelyn

Y.A. McKenna, Tom Postmes, and Ulf-Dietrich Reips (editors). *Oxford handbook of Internet psychology*. New York: Oxford University Press.
doi: https://doi.org/10.1093/oxfordhb/9780199561803.013.0016, accessed 20 June 2021.

Adam N Joinson and Carina B Paine, 2007. "Self-disclosure, privacy and the Internet," In: Adam N. Joinson, Katelyn Y.A. McKenna, Tom Postmes, and Ulf-Dietrich Reips (editors). *Oxford handbook of Internet psychology*. New York: Oxford University Press.

Hanna Krasnova, Natasha F Veltri, and Oliver Gunther, 2012. "Self-disclosure and privacy calculus on social networking sites: The role of culture," *Business & Information Systems Engineering*, volume 4, number 3, pp. 127–135.
doi: https://doi.org/10.1007/s12599-012-0216-6, accessed 20 June 2021.

Hanna Krasnova, Sarah Spiekermann, Ksenia Koroleva, and Thomas Hildebrand, 2010. "Online social networks: Why we disclose," *Journal of Information Technology*, volume 25, number 2, pp. 109–125.
doi: https://doi.org/10.1057/jit.2010.6, accessed 20 June 2021.

Robert S. Laufer and Maxine Wolfe, 1977. "Privacy as a concept and a social issue: A multidimensional developmental theory," *Journal of Social Issues*, volume 33, number 3, pp. 22–42.
doi: https://doi.org/10.1111/j.1540-4560.1977.tb01880.x, accessed 20 June 2021.

Richard S. Lazarus, 1991. *Emotion and adaptation*. New York: Oxford University Press.

Han Li, Rathindra Sarathy, and Heng Xu, 2010. "Understanding situational online information disclosure as a privacy calculus," *Journal of Computer Information Systems*, volume 51, number 1, pp. 62–71.

Han Li, Xin (Robert) Luo, Jie Zhang, and Heng Xu, 2017. "Resolving the privacy paradox: Toward a cognitive appraisal and emotion approach to online privacy behaviors," *Information & Management*, volume 54, number 8, pp. 1,012–1,022.
doi: https://doi.org/10.1016/j.im.2017.02.005, accessed 20 June 2021.

Liza Lin and Timothy Martin, 2020. "How coronavirus is eroding privacy," *Wall Street Journal* (15 April), at https://www.wsj.com/articles/coronavirus-paves-way-for-new-age-of-digital-surveillance-11586963028, accessed 20 June 2021.

Kun Liu and Evimaria Terzi, 2010. "A framework for computing the privacy scores of users in online social networks," *ACM Transactions on Knowledge Discovery from Data*, volume 5, number 1, article number 6, pp. 1–30.
doi: https://doi.org/10.1145/1870096.1870102, accessed 20 June 2021.

Mufan Luo and Jeffrey T Hancock, 2020. "Self-disclosure and social media: Motivations, mechanisms and psychological well-being," *Current Opinion in Psychology*, volume 31, pp. 110–115.
doi: https://doi.org/10.1016/j.copsyc.2019.08.019, accessed 20 June 2021.

Xiao Ma, Jeff Hancock, and Mor Naaman, 2016. "Anonymity, intimacy and self-disclosure in social media," *CHI '16: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pp. 3,857–3,869.
doi: https://doi.org/10.1145/2858036.2858414, accessed 20 June 2021.

Alice Marwick, Claire Fontaine, and danah boyd, 2017. "'Nobody sees it, nobody gets mad': Social media, privacy, and personal responsibility among low-SES youth," Social Media + Society (30 May).
doi: https://doi.org/10.1177/2056305117710455, accessed 20 June 2021.

Lorna McGregor, Daragh Murray, and Vivian Ng, 2018. "Four ways your Google searches and social media affect your opportunities in life," *The Conversation* (21 May), at https://theconversation.com/four-ways-your-google-searches-and-social-media-affect-your-opportunities-in-life-96809, accessed 20 June 2021.

Kamalich Muniz-Rodriguez, Sylvia K. Ofori, Lauren C. Bayliss, Jessica S. Schwind, Kadiatou Diallo, Manyun Liu, Jingjing Yin, Gerardo Chowell, and Isaac Chun-Hai Fung, 2020. "Social media use in emergency response to natural disasters: A systematic review with a public health perspective," *Disaster Medicine and Public Health Preparedness*, volume 14, number 1, pp. 139–149.
doi: https://doi.org/10.1017/dmp.2020.3, accessed 20 June 2021.

Teagen Nabity-Grover, Christy M.K. Cheung, and Jason Bennett Thatcher, 2020. "Inside out and outside in: How the

COVID-19 pandemic affects self-disclosure on social media," *International Journal of Information Management*, volume 55, 102188.
doi: https://doi.org/10.1016/j.ijinfomgt.2020.102188, accessed 20 June 2021.

Melanie Nguyen, Yu Sun Bin, and Andrew Campbell, 2012. "Comparing online and offline self-disclosure: A systematic review," *Cyberpsychology, Behavior, and Social Networking*, volume 15, number 2, pp. 103–111.
doi: https://doi.org/10.1089/cyber.2011.0277, accessed 20 June 2021.

PACT, 2020. "Private automated contact tracing," at https://pact.mit.edu, accessed 20 June 2021.

Leysia Palen, 2008. "Online social media in crisis events," *EDUCAUSE Quarterly*, volume 31, number 3, pp. 76–78, and at https://er.educause.edu/articles/2008/8/online-social-media-in-crisis-events, accessed 20 June 2021.

Joy V. Peluchette, Katherine Karl, Christa Wood, and Jennifer Williams, 2015. "Cyberbullying victimization: Do victims' personality and risky social network behaviors contribute to the problem?" *Computers in Human Behavior*, volume 52, pp. 424–435.
doi: https://doi.org/10.1016/j.chb.2015.06.028, accessed 20 June 2021.

Richard E. Petty, David DeSteno, and Derek D. Rucker, 2001. "The role of affect in attitude change," In: Joseph P. Forgas (editor). *Handbook of affect and social cognition*. Mahwah, N.J.: Lawrence Erlbaum Associates, pp. 212–233.

Betty Pfefferbaum and Carol S. North, 2020. "Mental health and the COVID-19 pandemic," *New England Journal of Medicine*, volume 383, number 6, pp. 510–512.
doi: https://doi.org/10.1056/NEJMp2008017, accessed 20 June 2021.

Ravi Philip Rajkumar, 2020. "COVID-19 and mental health: A review of the existing literature," *Asian Journal of Psychiatry*, volume 52, 102066.
doi: https://doi.org/10.1016/j.ajp.2020.102066, accessed 20 June 2021.

Koustuv Saha, John Torous, Eric D Caine, and Munmun De Choudhury, 2020. "Psychosocial effects of the COVID-19 pandemic: Large-scale quasi-experimental study on social media," *Journal of Medical Internet Research*, volume 22, number 11, e22600.
doi: https://doi.org/10.2196/22600, accessed 20 June 2021.

Antonia Sebastian, K.T. Lendering, B.L.M. Kothuis, A.D. Brand, Sebastiaan N. Jonkman, P.H.A.J.M. van Gelder, Maartje Godfroij, Bas Kolen, M. Comes, S.L.M. Lhermitte, Kenny Meesters, B.A. van de Walle, A. Ebrahimi Fard, S. Cunningham, N. Khakzad, and V. Nespeca, 2017. "Hurricane Harvey report: A fact-finding effort in the direct aftermath of hurricane Harvey in the greater Houston region," at https://repository.tudelft.nl/islandora/object/uuid%3A54c24519-c366-4f2f-a3b9-0807db26f69c, accessed 20 June 2021.

Deepa Seetharaman and Georgia Wells, 2017. "Hurricane Harvey victims turn to social media for assistance," *Wall Street Journal* (29 August), at https://www.wsj.com/articles/hurricane-harvey-victims-turn-to-social-media-for-assistance-1503999001, accessed 20 June 2021.

Kelly Servick, 2020. "Cellphone tracking could help stem the spread of coronavirus. Is privacy the price?" *Science* (22 March).
doi: https://doi.org/10.1126/science.abb8296, accessed 20 June 2021.

Carson Sievert and Kenneth Shirley, 2014. "LDAvis: A method for visualizing and interpreting topics," *Proceedings of the Workshop on Interactive Language Learning, Visualization, and Interfaces*, pp. 63–70.
doi: https://doi.org/10.3115/v1/W14-3110, accessed 20 June 2021.

H. Jeff Smith, Tamara Dinev, and Heng Xu, 2011. "Information privacy research: An interdisciplinary review," *MIS Quarterly*, volume 35, number 4, pp. 989–1,016.
doi: https://doi.org/10.2307/41409970, accessed 20 June 2021.

William Roth Smith, Keri K. Stephens, Brett R. Robertson, Jing Li, and Dhiraj Murthy, 2018. "Social media in citizen-led disaster response: Rescuer roles, coordination challenges, and untapped potential," *Proceedings of the International Conference on Information Systems for Crisis Response and Management (ISCRAM)*. pp. 639–648, and at https://par.nsf.gov/servlets/purl/10076203, accessed 20 June 2021.

Kate Starbird, Jim Maddock, Mania Orand, Peg Achterman, and Robert M. Mason, 2014. "Rumors, false flags, and digital vigilantes: Misinformation on Twitter after the 2013 Boston Marathon bombing," *iConference 2014 Proceedings*. doi: https://doi.org/10.9776/14308, accessed 20 June 2021.

Peter Stefanov, Kareem Darwish, Atanas Atanasov, and Preslav Nakov, 2020. "Predicting the topical stance and political leaning of media using tweets," *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pp. 527–537.
doi: https://doi.org/10.18653/v1/2020.acl-main.50, accessed 20 June 2021.

Yongqiang Sun, Nan Wang, Xiao-Liang Shen, and Jacky Xi Zhang, 2015. "Location information disclosure in location-based social network services: Privacy calculus, benefit structure, and gender differences," *Computers in Human Behavior*, volume 52, number C, pp. 278–292.
doi: https://doi.org/10.1016/j.chb.2015.06.006, accessed 20 June 2021.

Martin Szomszor, Patty Kostkova, and Connie St. Louis, 2011. "Twitter informatics: Tracking and understanding public reaction during the 2009 Swine Flu Pandemic," *2011 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology*, volume 1, pp. 320–323.
doi: https://doi.org/10.1109/WI-IAT.2011.311, accessed 20 June 2021.

Bruno Takahashi, Edson C. Tandoc, Jr., and Christine Carmichael, 2015. "Communicating on Twitter during a disaster: An analysis of tweets during Typhoon Haiyan in the Philippines," *Computers in Human Behavior*, volume 50, pp. 392–398.
doi: https://doi.org/10.1016/j.chb.2015.04.020, accessed 20 June 2021.

Lisa Collins Tidwell and Joseph B. Walther, 2002. "Computer-mediated communication effects on disclosure, impressions, and interpersonal evaluations: Getting to know one another a bit at a time," *Human Communication Research*, volume 28, number 3, pp. 317–348.
doi: https://doi.org/10.1111/j.1468-2958.2002.tb00811.x, accessed 20 June 2021.

Sabine Trepte, Leonard Reinecke, Nicole B. Ellison, Oliver Quiring, Mike Z. Yao, and Marc Ziegele, 2017. "A cross-cultural perspective on the privacy calculus," *Social Media + Society* (1 January).
doi: https://doi.org/10.1177/2056305116688035, accessed 20 June 2021.

Prasanna Umar, Anna Squicciarini, and Sarah Rajtmajer, 2019. "Detection and analysis of self-disclosure in online news commentaries," *WWW '19: The World Wide Web Conference*, pp. 3,272–3,278.
doi: https://doi.org/10.1145/3308558.3313669, accessed 20 June 2021.

Katja Valaskivi, Anna Rantasila, Mikihito Tanaka, and Risto Kunelius, 2019. *Traces of Fukushima: Global events, networked media and circulating emotions*. New York: Palgrave Pivot.
doi: https://doi.org/10.1007/978-981-13-6864-6, accessed 20 June 2021.

Asimina Vasalou, Alastair J. Gill, Fadhila Mazanderani, Chrysanthi Papoutsi, and Adam Joinson, 2011. "Privacy dictionary: A new resource for the automated content analysis of privacy," *Journal of the American Society for Information Science and Technology*, volume 62, number 11, pp. 2,095–2,105.
doi: https://doi.org/10.1002/asi.21610, accessed 20 June 2021.

Jessica Vitak and Jinyoung Kim, 2014. "'You can't block people offline': Examining how Facebooks affordances shape the disclosure process," *CSCW '14: Proceedings of the 17th ACM conference on Computer Supported Cooperative Work & Social Computing*, pp. 461–474.
doi: https://doi.org/10.1145/2531602.2531672, accessed 20 June 2021.

Yi-Chia Wang, Moira Burke, and Robert Kraut, 2016. "Modeling self-disclosure in social networking sites," *CSCW '16: Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, pp. 74–85.
doi: https://doi.org/10.1145/2818048.2820010, accessed 20 June 2021.

Gregor Wiedemann, Eugen Ruppert, and Chris Biemann, 2019. "UHH-LT at SemEval-2019 Task 6: Supervised vs. unsupervised transfer learning for offensive language detection," *Proceedings of the 13th International Workshop on Semantic Evaluation*, pp. 782–787.
doi: https://doi.org/10.18653/v1/S19-2137, accessed 20 June 2021.

Marc Wiedermann, E. Keith Smith, Jobst Heitzig, and Jonathan F. Donges, 2020. "A network-based microfoundation of

Granovetter's threshold model for social tipping," *Scientific Reports*, volume 10, article number 11202.
doi: https://doi.org/10.1038/s41598-020-67102-6, accessed 20 June 2021.

Yayouk E. Willems, Catrin Finkenauer, and Peter Kerkhof, 2020. "The role of disclosure in relationships," *Current Opinion in Psychology*, volume 31, pp. 33–37.
doi: https://doi.org/10.1016/j.copsyc.2019.07.032, accessed 20 June 2021.

Asta Zelenkauskaite, Nik Bessis, Stelios Sotiriadis, and Eleana Asimakopoulou, 2012. "Interconnectedness of complex systems of Internet of Things through social network analysis for disaster management," *2012 Fourth International Conference on Intelligent Networking and Collaborative Systems*, pp. 503–508.
doi: https://doi.org/10.1109/iNCoS.2012.25, accessed 20 June 2021.

Renwen Zhang, 2017. "The stress-buffering effect of self-disclosure on Facebook: An examination of stressful life events, social support, and mental health among college students," *Computers in Human Behavior*, volume 75, pp. 527–537.
doi: https://doi.org/10.1016/j.chb.2017.05.043, accessed 20 June 2021.

Renwen Zhang and Jiawei Sophia Fu, 2020. "Privacy management and self-disclosure on social network sites: The moderating effects of stress and gender," *Journal of Computer-Mediated Communication*, volume 25, number 3, pp. 236–251.
doi: https://doi.org/10.1093/jcmc/zmaa004, accessed 20 June 2021.

Xuan Zhao, Cliff Lampe, and Nicole B Ellison, 2016. "The social media ecology: User perceptions, strategies and challenges," *CHI '16: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pp. 89–100.
doi: https://doi.org/10.1145/2858036.2858333, accessed 20 June 2021.

Chen Zhao, Pamela Hinds, and Ge Gao, 2012. "How and to whom people share: The role of culture in self-disclosure in online communities," *CSCW '12: Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work*, pp. 67–76.
doi: https://doi.org/10.1145/2145204.2145219, accessed 20 June 2021.

Lichen Zhen, Yuanfeixue Nan, and Becky Pham, 2021. "College students coping with COVID-19: Stress-buffering effects of self-disclosure on social media and parental support," *Communication Research Reports*, volume 38, number 1, pp. 23–31.
doi: https://doi.org/10.1080/08824096.2020.1870445, accessed 20 June 2021.

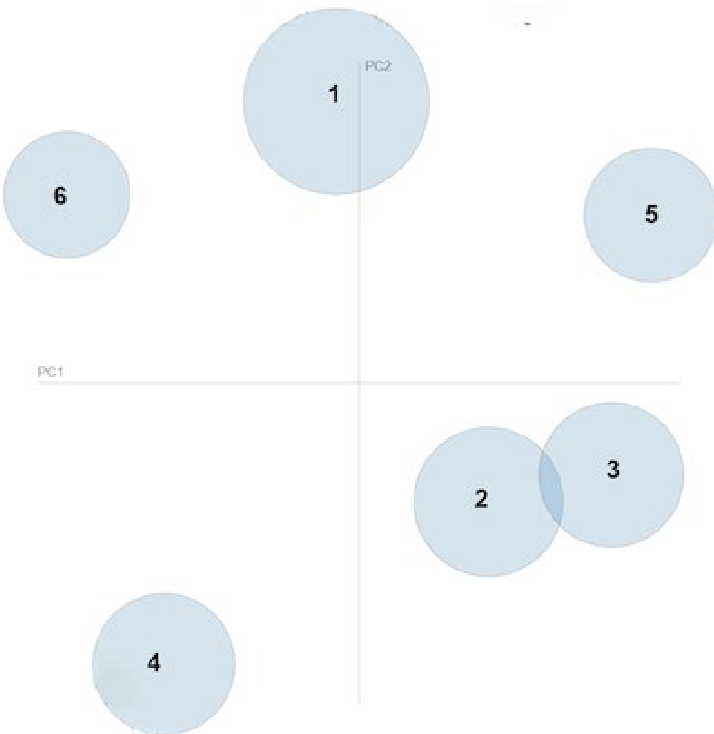
Appendix A

Table A1: Survey questions provided to MTurk workers for the labeling of each Tweet, adopted from Wang, *et al.* (2016).

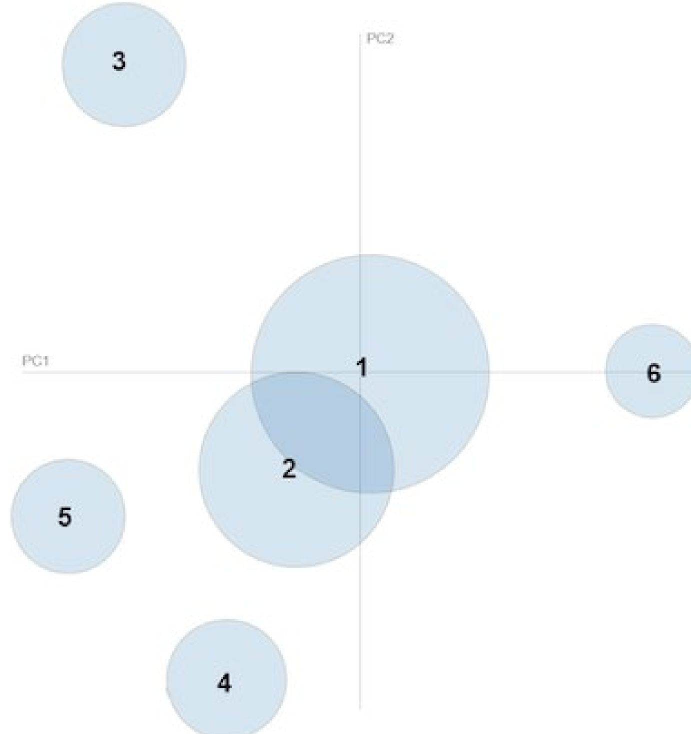
	To what extent does this post involve
A	personal information about the poster or people close to him/her, such as accomplishments, family, or problems the poster is having?
B	personal thoughts on past events, future plans, appearance, health, wishful ideas, etc.?
C	the poster's feelings and emotions, including concerns, frustrations, happiness, sadness, anger, and so on?
D	what is important to the poster in life?
E	the poster's close relationships with other people?

Table A2: Topical comparison of self-disclosing COVID-19 tweets between 16 May and 15 August 2020.	
Topic | Theme	**Top six keywords | Example**
1	get,day,love,time,year,feel
2	corona,death,covid,virus,die,get
3	china,trump,country,virus,president,world
4	covid19,health,test,need,community,help
5	home,stay,mask,safe,order,wear
6	pay,sir,money,june,help,exam
(a) 16 May — 15 June 2020	
Topic | Theme	**Top six keywords | Example**
1	home,get,like,stay,know,want
2	trump,american,president,vote,country
3	help,support,covid19,share,business,need
4	case,death,number,china,flu,virus
5	ear,distance,social,test,positive,school
6	play,student,cancel,exam,year,july
(b) 16 June — 15 July 2020	
Topic | Theme	**Top six keywords | Example**
1	get,like,home,want,stay,know
2	trump,distance,social,country,vote,american
3	lockdown,new,play,buy,game,watch
4	school,health,child,vaccine,risk,kid
5	money,pay,help,sir,sign,join
6	test,death,case,number,positive,flu
(b) 16 July — 15 August 2020	

Appendix B

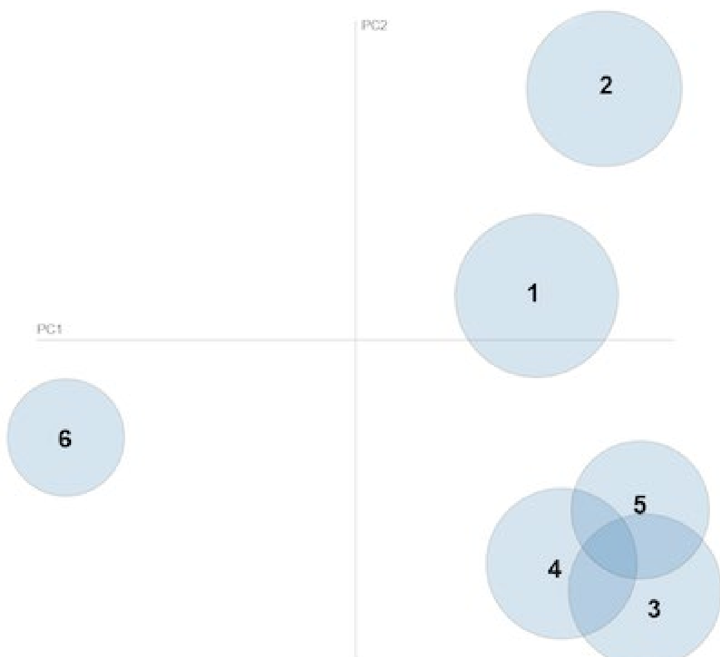


(a) Self-disclosing tweets

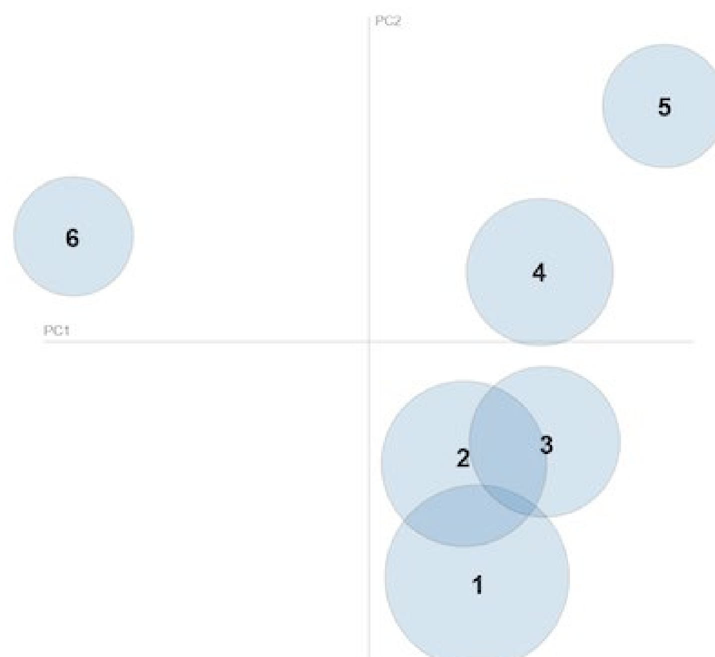


(b) Non-self-disclosing tweets

Figure B1: Intertopic distance maps for self-disclosing and non-self-disclosing COVID-19 tweets through 15 May 2020.

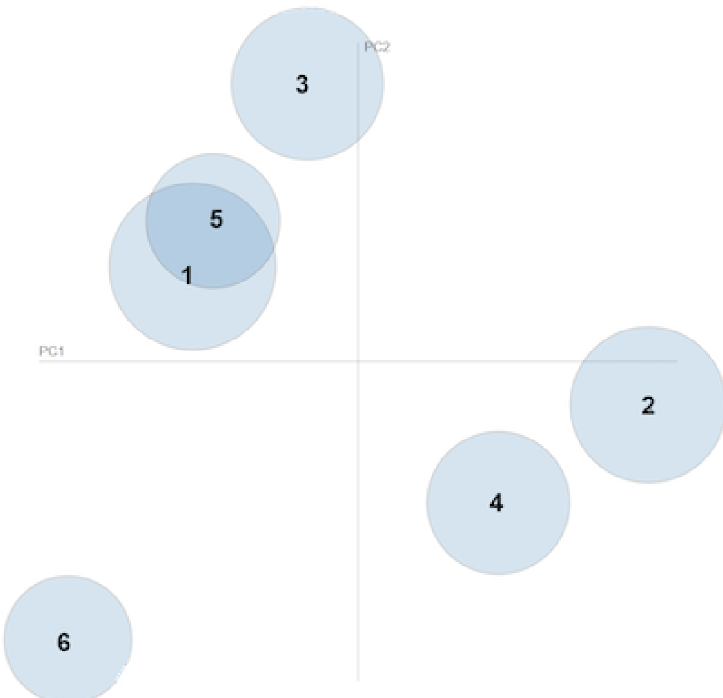


(a) Self-disclosing tweets

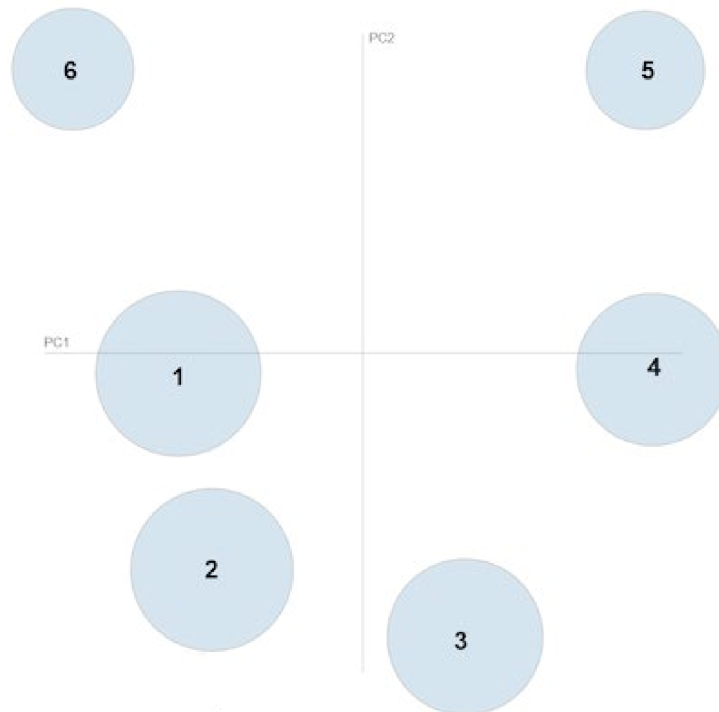


(b) Non-self-disclosing tweets

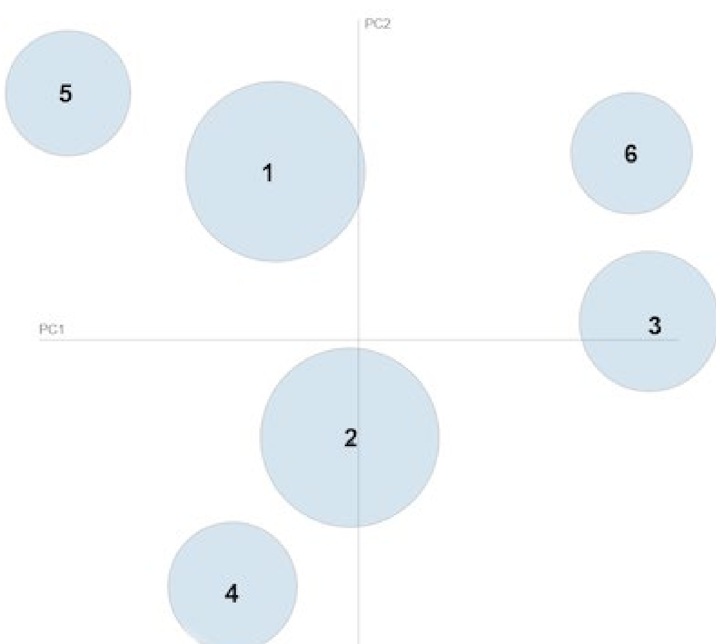
Figure B2: Intertopic distance maps for self-disclosing and non-self-disclosing tweets during Hurricane Harvey.



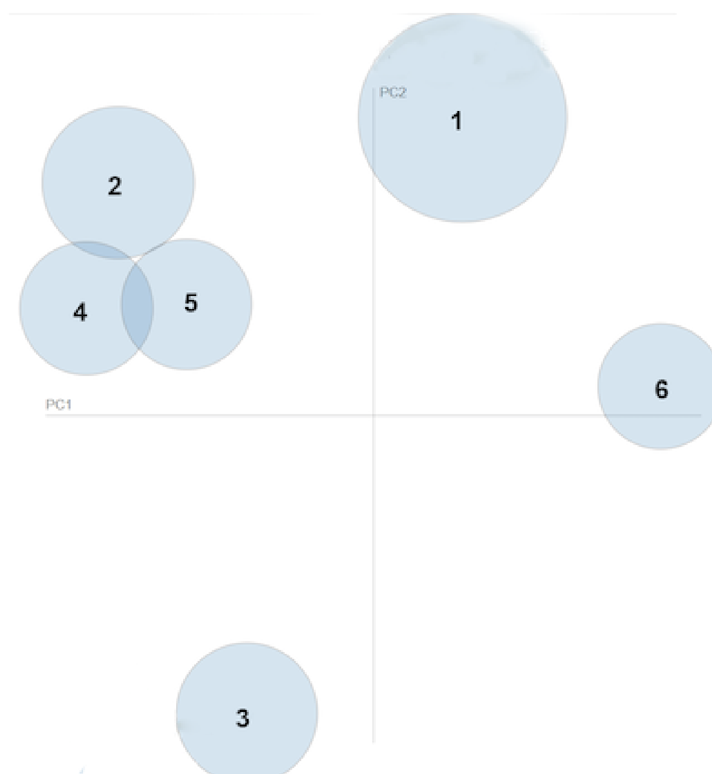
(a) 21 January — 15 March 2020



(b) 12 March — 15 May 2020



(c) 16 May — 15 June 2020



(d) 16 June — 15 July 2020

(e) 16 July — 15 August 2020

Figure B3: Intertopic distance maps for self-disclosing tweets in COVID-19 dataset.

Editorial history

Received 18 January 2021; revised 3 May 2021; accepted 21 June 2021.

A study of self-disclosure during the Coronavirus pandemic
by Taylor Blose, Prasanna Umar, Anna Squicciarini, and Sarah Rajtmajer.
First Monday, volume 26, number 7 (July 2021).
doi: http://dx.doi.org/10.5210/fm.v26i7.11555